Ahmedabad University

Started on	Monday, 20 December 2021, 5:33 PM
State	Finished
Completed on	Monday, 20 December 2021, 5:34 PM
Time taken	38 secs
Grade	**0** out of 2152 (**0**%)

Question 1

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

1710	✖	300	✖
1750	✖	700	✖
1400	✖	500	✖
790	✖	730	✖
2570	✖	900	✖
2320	✖	560	✖
1190	✖	280	✖
1510	✖	820	✖
1090	✖	990	✖
1280	✖	780	✖
1510	✖	790	✖
930	✖	730	✖
2360	✖	290	✖
890	✖	330	✖
2620	✖	900	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	233000	3	2	1710	4697	1.5	1710	300	1941
2	937000	3	1.75	2450	2691	2	1750	700	1915
3	667000	3	1	1400	1581	1.5	1400	500	1909
4	438000	3	1.75	1520	6380	1	790	730	1948
5	719000	4	2.5	2570	7173	2	2570	900	2005
6	580500	3	2.5	2320	3980	2	2320	560	2003
7	280000	2	1.5	1190	1265	3	1190	280	2005
8	687500	4	1.75	2330	5000	1.5	1510	820	1929
9	535000	3	1	1090	3000	1.5	1090	990	1929
10	322500	4	2.75	2060	6659	1	1280	780	1981
11	696000	3	2.5	2300	3060	1.5	1510	790	1930
12	550000	4	1	1660	34848	1	930	730	1933
13	640000	4	2	2360	6000	2	2360	290	1904
14	240000	4	1	1220	8075	1	890	330	1969
15	605000	4	2.5	2620	7553	2	2620	900	1996
16	625000	4	2.5	2570	5520	2	2570	1500	2000
17	775000	4	2.25	4220	24186	1	2600	1620	1984
18	861990	5	2.75	3595	5639	2	3595	400	2014
19	685000	3	1	1570	2280	2	1570	600	1922
20	309000	3	1	1280	9656	1	920	360	1959

21	488000	3	2.5	3160	13603	2	3160	600	2003
22	210490	3	1	990	8528	1	990	550	1966
23	785000	4	2.5	2290	13416	2	2290	750	1981
24	450000	3	1.75	1250	5963	1	1250	890	1953
25	1350000	3	2.5	2753	65005	1	2165	588	1953

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**

2. Bathroom ✖ **(2 Marks)**

3. Sqft_Living ✖ **(2 Marks)**

4. Sqft_lot ✖ **(2 Marks)**

5. Sqft_above ✖ **(2 Marks)**

6. Sqft_basement ✖ **(2 Marks)**

7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**

2. p-value slope for Bedroom = ✖ **(1 Mark)**

3. p-value slope for Bathroom = ✖ **(1 Mark)**

4. p-value slope for Sqft_Living = ✖ **(1 Mark)**

5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 65.5_% of variance in dependent variable is explained by independent variables
- 67.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 65.5_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.

- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 2

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	2620	✖	567	✖
2	1670	✖	1300	✖
3	1050	✖	899	✖
4	1280	✖	900	✖
5	1030	✖	578	✖
6	1390	✖	510	✖
7	1740	✖	1330	✖
8	1760	✖	890	✖
9	930	✖	678	✖
10	880	✖	789	✖
11	1310	✖	455	✖
12	1560	✖	570	✖
13	1120	✖	360	✖
14	1500	✖	344	✖
15	1670	✖	555	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	648000	4	2.5	2620	5450	2	2620	567	2001
2	469500	5	2.5	2970	24759	1	1670	1300	1969
3	237500	3	1.75	1050	7854	1	1050	899	1975
4	216000	3	1	1280	8712	1	1280	900	1956
5	427200	3	1	1030	8400	1	1030	578	1963
6	280000	3	2.25	1900	7800	1	1390	510	1977
7	718000	5	3	3070	9804	1	1740	1330	1968
8	420000	2	2.75	1760	4139	2	1760	890	2010
9	379000	2	1	930	5400	1	930	678	1952
10	326100	2	1	880	7683	1	880	789	1942
11	535000	3	2.5	1360	1016	3	1310	455	2003
12	283000	4	2.75	2130	8560	1	1560	570	1992
13	260000	3	1	1480	7469	1.5	1120	360	1940
14	399000	4	1	1500	6388	1.5	1500	344	1951
15	484950	3	2.25	1670	5004	2	1670	555	1987
16	640000	4	1.75	2060	9828	1	2060	1000	1960
17	585000	4	2.5	2160	8158	1	1660	500	1952
18	645000	4	2	2780	11583	1	1190	1590	1955
19	253000	3	2	1490	7651	1	1490	310	1988
20	178500	2	1	740	6460	1	740	788	1953
21	291000	4	2.5	1860	6325	2	1860	120	1991
22	199950	2	2.75	1590	20917	1.5	1590	666	1920
23	553500	2	1	850	2340	1	850	700	1922
24	189950	2	1	1030	4188	1	1030	900	1981
25	289000	3	2	1850	9550	1	1850	455	1988

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ (1 Mark)

7. p-value slope for sqft_above = ✖ (1 Mark)

8. p-value slope for sqft_basement = ✖ (1 Mark)

9. p-value slope for yr_built = ✖ (1 Mark)

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 75.42_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 75.42_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 3

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1410	✖	455	✖
2	1330	✖	699	✖
3	1150	✖	340	✖
4	2020	✖	788	✖
5	3130	✖	900	✖
6	2670	✖	345	✖
7	1980	✖	677	✖
8	2100	✖	1040	✖
9	1350	✖	910	✖
10	1190	✖	950	✖
11	2208	✖	788	✖
12	1010	✖	270	✖
13	970	✖	610	✖
14	1520	✖	1830	✖
15	2740	✖	900	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	435000	5	1	1410	6750	1.5	1410	455	1929
2	601000	3	1.75	1330	6743	1	1330	699	1958
3	460500	3	1	1490	10650	1	1150	340	1963
4	720000	3	2.5	2020	1159	3	2020	788	2000
5	851000	4	2.5	3130	13202	2	3130	900	1996
6	550000	4	3	2670	5000	2	2670	345	1916
7	700000	4	1	1980	4560	1.5	1980	677	1920
8	323000	3	1.75	2100	14850	1	2100	1040	1963
9	345500	3	1	1350	8581	1	1350	910	1944
10	378000	4	1.5	2140	7920	1	1190	950	1959
11	650000	4	2	2208	5000	3	2208	788	1917
12	385000	3	2	1010	7380	1	1010	270	1982
13	425000	2	1	970	2970	1	970	610	1910
14	210000	4	2	1520	6174	1.5	1520	1830	1920
15	630000	4	2.5	2740	43101	2	2740	900	1993
16	335000	3	1.75	1570	7500	1	1300	270	1953
17	196500	3	1	1320	9000	1	1320	677	1966
18	582800	4	2.75	2550	7636	1	1440	1110	1986
19	440000	3	1.75	1560	7207	1	1250	310	1983
20	752500	4	2.25	2360	8616	2	2360	788	1974
21	532500	3	1.75	1330	5000	2	1210	120	1909
22	265000	3	1	1070	9000	1	1070	666	1950
23	397500	3	1.25	1510	13737	1	810	700	1961
24	524250	3	2.5	1710	3469	2	1710	900	1999
25	765000	4	2.75	2790	10819	2	2790	345	1996

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 70.62_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 70.62_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 4

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr. No.	sqft_above		sqft_basement	
1	2340	✖	630	✖
2	2570	✖	90	✖
3	2560	✖	1250	✖
4	950	✖	346	✖
5	1080	✖	400	✖
6	1250	✖	800	✖
7	1654	✖	999	✖
8	1410	✖	1100	✖
9	1500	✖	660	✖
10	700	✖	700	✖
11	1420	✖	677	✖
12	1600	✖	900	✖
13	1410	✖	555	✖
14	1050	✖	1050	✖
15	1750	✖	566	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	905000	4	3.5	2970	14486	2	2340	630	1997
2	800000	4	2.5	2570	50308	1.5	2570	90	1993
3	839900	4	3.5	3810	13592	1	2560	1250	2013
4	200000	1	1.5	1010	1157	2	950	346	2007
5	300000	3	2	1480	6698	1	1080	400	1979
6	530000	3	1.75	1250	6041	1.5	1250	800	1942
7	314950	3	2.25	1654	8479	2	1654	999	1995
8	255000	3	1.75	1410	9315	1	1410	1100	1960
9	550000	5	2.75	2160	5720	1	1500	660	1950
10	435000	3	1	1400	4800	1	700	700	1917
11	234000	3	1.75	1420	8738	1	1420	677	1966
12	615000	4	2.25	2500	10062	1	1600	900	1975
13	205000	3	2	1410	8384	1	1410	555	1985
14	384000	4	1.75	2100	7135	1	1050	1050	1955
15	324000	3	2.5	1750	7208	2	1750	566	1994
16	645000	3	1.75	2070	5500	1	1130	940	1946
17	570000	3	3.25	1570	1777	2	1260	310	2007
18	324000	3	2.5	1920	5322	2	1920	210	1999
19	270000	2	1	1000	10320	1	1000	600	1943
20	290000	5	1.5	2120	7700	1.5	2120	1380	1962
21	250000	3	2	1440	9220	1	1440	800	1965
22	342000	3	1.75	1660	16275	2	1660	350	1990
23	255000	3	1.5	1340	8450	1	1340	340	1958
24	800000	3	2.25	2380	11824	1	1450	930	1972
25	451000	3	1.75	1560	4049	1.5	1000	560	1926

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 85_% of variance in dependent variable is explained by independent variables
- 89.19_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 89.19_ % of variance in dependent variable is explained by independent variable

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 5

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	3064	✖	455	✖
2	1530	✖	677	✖
3	1410	✖	500	✖
4	1020	✖	900	✖
5	1150	✖	347	✖
6	870	✖	390	✖
7	920	✖	1000	✖
8	1800	✖	444	✖
9	1360	✖	680	✖
10	1430	✖	770	✖
11	2160	✖	1040	✖
12	1780	✖	620	✖
13	1140	✖	123	✖
14	1330	✖	1180	✖
15	1940	✖	690	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	1095000	0	0	3064	4764	3.5	3064	455	1990
2	147500	3	1	1530	8498	1	1530	677	1959
3	390000	3	2	1910	11576	1	1410	500	1978
4	229000	3	1	1020	8100	1	1020	900	1954
5	185000	3	1	1150	8100	1	1150	347	1932
6	414900	3	1.5	1260	9570	1	870	390	1941
7	200000	4	2	1920	4822	1	920	1000	1914
8	225000	4	2.25	1800	9350	1	1800	444	1969
9	720000	3	1.75	2040	4000	2	1360	680	1924
10	930000	4	2.5	2200	4000	2	1430	770	1908
11	1270000	5	2.5	3200	17204	1	2160	1040	1952
12	700000	5	3.25	2400	3118	2	1780	620	1928
13	466000	2	1.5	1140	1058	3	1140	123	2005
14	374000	3	1.5	1330	10640	1	1330	1180	1976
15	480000	3	2.5	1940	10035	2	1940	690	1994
16	740000	3	2.5	3000	25341	2	3000	456	1995
17	1010000	4	1	1820	5400	1.5	1820	555	1923
18	395000	4	2.5	2910	5000	2	2910	1220	2002
19	340000	2	1	1090	6771	1	1090	420	1954
20	209000	3	2	1230	1340	2	1020	210	2003
21	431650	5	2.5	1710	7700	1.5	1710	600	1962
22	980000	5	2.5	3160	11470	1	1780	1380	1971
23	660000	3	2.5	2680	28243	2	2680	800	1993
24	730000	4	2.25	2190	9009	2	1840	350	1977
25	291500	3	1.75	1260	9600	1	1260	340	1988

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 75_% of variance in dependent variable is explained by independent variables
- 76.96_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 76.96_ % of variance in dependent variable is explained by independent variable

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 6

Not answered

Marked out of 7

A sample of 32 money-market mutual funds was chosen on January 1, 2016, and the average annual rate of return over the past 30 days was found to be 3.23 percent, and the sample standard deviation was 0.51 percent. A year earlier, a sample of 38 money-market funds showed an average rate of return of 4.36 percent, and the sample standard deviation was 0.84 percent.

Calculated Value of Z = ✖

The correct answer is: -6.9155

Question 7

Not answered

Marked out of 7

Two research laboratories have independently produced drugs that provide relief to arthritis sufferers. The first drug was tested on a group of 90 arthritis sufferers and produced an average of 8.5 hours of relief, and a sample standard deviation of 1.8 hours. The second drug was tested on 80 arthritis sufferers, producing an average of 7.9 hours of relief, and a sample standard deviation of 2.1 hours. At the 0.05 level of significance, the calculated value of z = ✖

The correct answer is: 1.99

Question 8

Not answered

Marked out of 7

Jeff Richardson, the receiving clerk for a chemical-products distributor, is faced with the continuing problem of broken glassware, including test-tubes, petri dishes, and flasks. Jeff has determined some additional shipping precautions that can be undertaken to prevent breakage, and he has asked the Purchasing Director to inform the suppliers of the new measures. Data for 8 suppliers are given below in terms of average number of broken items per shipment. ✖ is the calculated value of t, at 5% significant level , that the new measures have lowered the average number of broken items?

Supplier	1	2	3	4	5	6	7	8
Before	16	12	18	7	14	19	6	17
After	14	13	12	6	9	15	8	15

The correct answer is: -2.147

Question 9

Not answered

Marked out of 7

General Electric has developed a new bulb whose design specifications call for a light output of 960 lumens compared to an earlier model that produced only 750 lumens. The company's data indicate that the standard deviation of light output for this type of bulb is 18.4 lumens. From a sample of 20 new bulbs, the testing committee found an average light output of 954 lumens per bulb. At a 0.05 significance level, ✖ is the value of z such that General Electric conclude that its new bulb is producing the specified 960 lumens output?

The correct answer is: -1.47

Question 10

Not answered

Marked out of 7

Some financial theoreticians believe that the stock market's daily prices constitute a "random walk with positive drift". If this is accurate, then the Dow Jones Industrial Average should show a gain on more than 50 percent of all trading days. If the average increased on 101 of 175 randomly chosen days, ✖ is the calculated value of z? Use a 0.01 level of significance.

The correct answer is: -2.04

Question 11

Not answered

Marked out of 7

In a sample of 19 observations from a normal distribution with mean 18 and standard deviation 4.8. P(16≤ x̄≤ 20) is ✖.

The correct answer is: 0.9312

Question 12

Not answered

Marked out of 7

From a population of 75 items with a mean of 364 and a variance of 18, 32 items were randomly selected without replacement.

The standard error of the mean is ✖.

The correct answer is: 0.5713

Question 13

Not answered

Marked out of 7

In an automotive safety test conducted by the New Delhi Highway Safety Research Center, the average tire pressure in a sample of 62 tires was found to be 24 pounds per square inch, and the standard deviation was 2.1 pounds per square inch.

The lower limit and upper limit of 95% confidence interval for the population mean are ✖ and _____ respectively.

The correct answer is: 23.48,24.52

Question 14

Not answered

Marked out of 7

Because refunds are paid more quickly on tax returns that are filed electronically, the Commissioner of the Internal Revenue Service was wondering whether refunds due on returns filed by mail were smaller than those due on returns filed electronically. Looking only at returns claiming refunds, a sample of 17 filed by mail had an average refund of $563, and a standard deviation of $378. The average refund on a sample of 13 electronically filed returns was $958, and the sample standard deviation was $619. At 0.01 significant level, calculated value of t = ✖.

The correct answer is: -2.1624

Question 15

Not answered

Marked out of 7

A television documentary on overeating claimed that Americans are about 10 pounds over-weight on average. To test this claim, eighteen randomly selected individuals were examined; their average excess weight was found to be 12.4 pounds, and the sample standard deviation was 2.7 pounds. At a significance level of 0.01.

The calculated value of t = ✖.

The correct answer is: 3.7712

Question 16

Not answered

Marked out of 7

Each day, the United States Customs Service has historically intercepted about $28 million in contraband goods being smuggled into the country, with a standard deviation of $16 million per day. On 64 randomly chosen days in 1992, the U.S. Customs Service intercepted an average of $30.3 million in contraband goods.

The calculated value of z = ✖ (consider 5% significant level)

The correct answer is: 1.15

Question 17

Not answered

Marked out of 7

A large stock-brokerage firm wants to determine how successful its new account executives have been at recruiting clients. After completing their training, new account execs spend several weeks calling prospective clients, trying to get the prospects to open accounts with the firm. The following data give the numbers of new accounts opened in their first 2 weeks by 10 randomly chosen female account execs and by 8 randomly chosen male account execs.

Number of New Accounts										
Female Account execs	12	11	14	13	13	14	13	12	14	12
Male account execs	13	10	11	12	13	12	10	12		

The calculated value of t = ✖ (at 5% level of significance)

The correct answer is: 2.2454

Question 18

Not answered

Marked out of 7

MacroSwift estimated last year that 35 percent of potential software buyers were planning to wait to purchase the new operating system, Window Panes, until an upgrade had been released. After an advertising campaign to reassure the public, MacroSwift surveyed 3,000 people and found 950 who were still sceptical. At the 5% significance level, the calculated value of z = ✖ ?

The correct answer is: -3.8276

Question 19

Not answered

Marked out of 7

Additives-R-Us has developed an additive to improve fuel efficiency for trucks that pull very heavy loads. They tested the additive by randomly selecting 18 trucks and dividing them into 9 pairs. In each pair, both trucks hauled the same type of load over the same roadway, but only one truck used fuel with the new additive. Different pairs followed different routes and carried different loads. The resulting fuel efficiencies (in miles per gallon) are given below. At 0.01 significant level, ✖ is the calculated value of t, such that the trucks using fuel with the additive achieved significantly better fuel efficiency than trucks using regular fuel?

Pair	1	2	3	4	5	6	7	8	9
Regular	5.7	6.1	5.9	6.2	6.4	5.1	5.9	6.0	5.5
Additive	6.0	6.2	5.8	6.6	6.7	5.3	5.7	6.1	5.9

The correct answer is: 2.358

Question 20

Not answered

Marked out of 7

Rick Douglas, the new manager of Food Barn, is interested in the percentage of customers who are totally satisfied with the store. The previous manager had 86 percent of the customers totally satisfied, and Rick claims the same is true today. Rick sampled 187 customers and found 157 were totally satisfied. At the 1% significance level, the calculated value of z = ✖

The correct answer is: -0.80

Question 21

Not answered

Marked out of 7

Realtor Elaine Snyderman took a random sample of 12 homes in a prestigious suburb of Chicago and found the average appraised market value to be $780,000 and the standard deviation was $49,000. ✖ is the calculated value of z for all homes in the area, the mean appraised value is $825,000 against the alternative that it is less than $825,000. Use the 0.05 level of significance.

The correct answer is: -3.181

Question 22

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1900	✖	677	✖
2	1050	✖	900	✖
3	1180	✖	340	✖
4	1750	✖	640	✖
5	2400	✖	345	✖
6	1790	✖	555	✖
7	1720	✖	123	✖
8	1800	✖	1180	✖
9	1520	✖	690	✖
10	1150	✖	456	✖
11	380	✖	555	✖
12	910	✖	1220	✖
13	1230	✖	420	✖
14	2480	✖	400	✖
15	2050	✖	780	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	535800	4	2.5	1900	5790	2	1900	677	1994
2	287653	3	1	1050	16050	1	1050	900	1960
3	438000	3	2.5	1520	1304	2	1180	340	2006
4	450000	4	1.75	2390	23899	1	1750	640	1949
5	327000	5	2.75	2400	8050	2	2400	345	1998
6	370500	4	2.5	1790	6120	2	1790	555	2005
7	265000	3	2.5	1720	6271	2	1720	123	1993
8	450000	4	2.75	2980	13260	1	1800	1180	1979
9	425000	3	1	1520	213444	1.5	1520	690	1988
10	166600	3	1.75	1150	8690	1	1150	456	1977
11	245000	1	0.75	380	15000	1	380	555	1963
12	565000	3	1	910	5212	1	910	1220	1951
13	539900	3	1.75	1650	10150	1	1230	420	1976
14	910000	3	3.5	2480	3200	2	2480	400	2010
15	280000	4	2.5	2050	7416	2	2050	780	1990
16	275000	3	2.25	1420	8549	2	1420	345	1989
17	467000	3	2.5	3460	6590	2	3460	456	2001
18	457500	3	2.5	1500	4445	2	1500	330	1996
19	381156	2	1	1320	3090	1	1320	234	1908
20	289999	5	2.5	2180	8240	1	1220	960	1977
21	642450	3	1.75	1830	4160	1	1230	600	1919
22	469950	3	2.25	1620	8701	1	1220	400	1962
23	795000	3	2	2920	13650	1	1460	1460	1975
24	785000	4	2.5	3300	10514	2	3300	456	1984
25	860000	3	2.25	3060	12095	2	3060	700	1983

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 55_% of variance in dependent variable is explained by independent variables
- 56.58_ % of variance in dependent variable is explained by independent variable
- 50.12_% of variance in independent variables are explained by dependent variables
- 52.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 56.58_ % of variance in dependent variable is explained by independent variable

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 23

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	970	✖	455	✖
2	1510	✖	500	✖
3	1300	✖	677	✖
4	1260	✖	900	✖
5	2610	✖	345	✖
6	1110	✖	800	✖
7	1780	✖	678	✖
8	1900	✖	543	✖
9	1340	✖	1110	✖
10	1380	✖	673	✖
11	1230	✖	690	✖
12	1350	✖	456	✖
13	1350	✖	555	✖
14	1290	✖	1220	✖
15	2470	✖	789	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	226000	3	1	970	5000	1	970	455	1968
2	825000	3	1.75	2010	3090	1.5	1510	500	1926
3	175000	4	1	1300	6030	1.5	1300	677	1947
4	465750	3	1.5	1260	10350	1	1260	900	1959
5	659000	4	2.25	2610	24931	2	2610	345	1983
6	595000	3	1.75	1910	5753	1	1110	800	1941
7	350000	3	2.25	1780	16290	2	1780	678	1987
8	270000	4	2.25	1900	8600	1	1900	543	1975
9	662500	6	2.25	2450	25600	1	1340	1110	1957
10	249900	3	1.75	1380	14000	1	1380	673	1939
11	324500	3	1.75	1920	11340	1	1230	690	1977
12	430000	2	1.75	1350	4003	1	1350	456	2004
13	419950	2	1.75	1350	4003	1	1350	555	2004
14	285000	4	1.75	2510	7440	1	1290	1220	1962
15	332000	4	2.5	2470	7780	2	2470	789	2002
16	282000	4	1.75	1660	10725	1	960	700	1956
17	320000	3	1	1340	5175	1	940	400	1987
18	428750	3	1	1620	30736	1.5	1620	780	1911
19	330000	3	2.25	2300	35287	2	2300	345	1977
20	456500	4	3.5	2000	2309	3	2000	456	2008
21	352000	3	2.25	1640	11050	1	1640	330	1972
22	664000	4	2	1530	4080	1.5	1530	234	1912
23	628000	3	1.75	4000	11894	1	2190	1810	1987
24	512000	3	1.75	1610	12555	1	1080	530	1977
25	254000	3	2	2070	9000	1	1450	620	1969

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 35.96_% of variance in dependent variable is explained by independent variables
- 36.96_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 35.96_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 24

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1340	✖	899	✖
2	1680	✖	1510	✖
3	1630	✖	900	✖
4	2990	✖	345	✖
5	1550	✖	789	✖
6	1690	✖	765	✖
7	1960	✖	456	✖
8	1360	✖	1030	✖
9	2600	✖	890	✖
10	1950	✖	988	✖
11	1250	✖	1000	✖
12	840	✖	340	✖
13	1680	✖	900	✖
14	2220	✖	345	✖
15	3090	✖	1800	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	650000	3	2	1340	2720	1.5	1340	899	1913
2	1216000	4	2.5	3190	8684	1	1680	1510	1967
3	422120	3	2.5	1630	4534	2	1630	900	1987
4	625000	5	2.5	2990	15085	2	2990	345	2007
5	385200	4	1	1550	7740	1.5	1550	789	1954
6	500000	3	2.25	1690	4964	2	1690	765	1989
7	308000	4	2.25	1960	12243	2	1960	456	1989
8	565000	3	2.75	2390	9966	1	1360	1030	1977
9	581000	3	2.5	2600	4438	2	2600	890	2004
10	361000	3	2	1950	8698	1	1950	988	1999
11	345000	3	1	1250	7210	1	1250	1000	1964
12	419000	2	2.25	1180	1253	2	840	340	2008
13	347000	3	2.5	1680	4308	2	1680	900	2004
14	625000	4	2.25	2220	36085	2	2220	345	1989
15	2400000	4	4.25	4890	15188	2	3090	1800	1999
16	379000	3	3.25	2660	17852	2.5	2660	890	2014
17	310000	3	1.5	1150	3323	2	1150	780	1988
18	250000	2	2	1130	5500	1	1130	345	1941
19	682000	5	2.75	2760	5000	2	2760	456	2005
20	246500	3	2	1430	8919	1	1430	330	1992
21	325000	3	1	1490	57381	1.5	1490	234	1932
22	310950	4	2.5	2030	4997	2	2030	1290	2004
23	1100000	3	2	3010	5000	2	1890	1120	1931
24	287500	3	1	1810	7200	1	1130	680	1954
25	431000	3	1	1810	7200	1	1130	680	1954

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 86.52_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 85.12_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 86.52_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 25

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1340	✖	899	✖
2	1680	✖	1510	✖
3	1630	✖	900	✖
4	2990	✖	345	✖
5	1550	✖	789	✖
6	1690	✖	765	✖
7	1960	✖	456	✖
8	1360	✖	1030	✖
9	2600	✖	890	✖
10	1950	✖	988	✖
11	1250	✖	1000	✖
12	840	✖	340	✖
13	1680	✖	900	✖
14	2220	✖	345	✖
15	3090	✖	1800	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	650000	3	2	1340	2720	1.5	1340	899	1913
2	1216000	4	2.5	3190	8684	1	1680	1510	1967
3	422120	3	2.5	1630	4534	2	1630	900	1987
4	625000	5	2.5	2990	15085	2	2990	345	2007
5	385200	4	1	1550	7740	1.5	1550	789	1954
6	500000	3	2.25	1690	4964	2	1690	765	1989
7	308000	4	2.25	1960	12243	2	1960	456	1989
8	565000	3	2.75	2390	9966	1	1360	1030	1977
9	581000	3	2.5	2600	4438	2	2600	890	2004
10	361000	3	2	1950	8698	1	1950	988	1999
11	345000	3	1	1250	7210	1	1250	1000	1964
12	419000	2	2.25	1180	1253	2	840	340	2008
13	347000	3	2.5	1680	4308	2	1680	900	2004
14	625000	4	2.25	2220	36085	2	2220	345	1989
15	2400000	4	4.25	4890	15188	2	3090	1800	1999
16	379000	3	3.25	2660	17852	2.5	2660	890	2014
17	310000	3	1.5	1150	3323	2	1150	780	1988
18	250000	2	2	1130	5500	1	1130	345	1941
19	682000	5	2.75	2760	5000	2	2760	456	2005
20	246500	3	2	1430	8919	1	1430	330	1992
21	325000	3	1	1490	57381	1.5	1490	234	1932
22	310950	4	2.5	2030	4997	2	2030	1290	2004
23	1100000	3	2	3010	5000	2	1890	1120	1931
24	287500	3	1	1810	7200	1	1130	680	1954
25	431000	3	1	1810	7200	1	1130	680	1954

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 86.52_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 85.12_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 86.52_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 26

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1000	✖	1000	✖
2	1730	✖	160	✖
3	1060	✖	850	✖
4	2710	✖	876	✖
5	3100	✖	1760	✖
6	1250	✖	1000	✖
7	970	✖	620	✖
8	1050	✖	190	✖
9	2000	✖	555	✖
10	3610	✖	140	✖
11	1930	✖	567	✖
12	1280	✖	890	✖
13	2090	✖	780	✖
14	1140	✖	345	✖
15	1800	✖	456	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	420000	3	2.25	2000	8030	1	1000	1000	1963
2	269950	3	2.5	1890	4838	2	1730	160	2002
3	493000	4	2	1910	2874	1	1060	850	1910
4	370000	4	2.5	2710	5880	2	2710	876	1998
5	2250000	5	4.25	4860	9453	1.5	3100	1760	1905
6	305000	3	1	1250	23680	1	1250	1000	1967
7	400000	3	1.75	1590	8219	1.5	970	620	1938
8	320000	3	1.5	1240	1221	2	1050	190	2009
9	715000	4	1	2000	4800	2	2000	555	1911
10	650000	6	2.75	3610	10003	1.5	3610	140	1966
11	218000	5	1.75	1930	8040	1	1930	567	1966
12	500000	2	1	1280	5400	1	1280	890	1964
13	460000	4	2	2090	40419	1	2090	780	1984
14	232500	3	1.5	1940	9887	1	1140	345	1969
15	408200	3	2.5	1800	5761	2	1800	456	1990
16	522000	3	1.75	1730	8400	1	1400	330	1980
17	675000	4	2.5	2810	11120	2	2810	234	1982
18	1355000	4	3.5	3550	11000	1	2260	1290	1999
19	249950	4	1	1330	7980	1.5	1330	890	1952
20	210000	2	1	750	34133	1	750	789	1950
21	429950	4	1	1760	7216	1	1090	670	1947
22	737500	4	2.5	3200	36276	2	3200	899	1993
23	710000	4	2	1740	9000	1	1740	900	1958
24	470000	3	2.5	2100	4700	2	2100	234	1999
25	1100000	4	3.5	4270	40097	1	4270	450	1993

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 80.35_% of variance in dependent variable is explained by independent variables
- 97.89_ % of variance in dependent variable is explained by independent variable
- 87.01_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 87.01_% of variance in independent variables are explained by dependent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 27

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	800	✖	350	✖
2	1140	✖	1010	✖
3	1230	✖	140	✖
4	2380	✖	345	✖
5	770	✖	600	✖
6	1070	✖	530	✖
7	2030	✖	1720	✖
8	2910	✖	1040	✖
9	2400	✖	678	✖
10	2830	✖	900	✖
11	2440	✖	555	✖
12	1050	✖	140	✖
13	1400	✖	567	✖
14	2730	✖	890	✖
15	1610	✖	780	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	329999	3	1150	6908	1	800	350	1952
2	950000	3	2150	4200	1	1140	1010	1960
3	447000	3	1370	6001	1	1230	140	1944
4	699800	2	2380	6600	1	2380	345	2010
5	850000	3	1370	3850	1	770	600	1911
6	316500	3	1600	14250	1	1070	530	1979
7	942500	5	3750	9612	1	2030	1720	1981
8	1310000	2	3950	3938	2	2910	1040	1991
9	856600	4	2400	13430	1	2400	678	1964
10	405600	5	2830	4000	2.5	2830	900	1918
11	565000	4	2440	22594	2	2440	555	1996
12	400000	3	1190	1149	2	1050	140	2008
13	330000	3	1400	218252	1	1400	567	1997
14	374950	4	2730	4683	2	2730	890	2005
15	405300	3	2390	7939	1	1610	780	1983
16	560000	3	2060	31400	2	2060	345	1984
17	824000	4	2490	9864	1	1290	1200	1961
18	260000	3	1100	10968	1	1100	899	1984
19	760000	4	2430	6099	1	1470	960	1965
20	405000	3	3170	12750	2	2360	810	1995
21	429900	4	2640	8625	2	2640	640	1987
22	341500	3	1130	7223	1	1130	660	1961
23	465250	5	1940	7642	1.5	1940	1110	1957
24	450000	2	1100	3000	1.5	1100	444	1912
25	390000	3	1200	10800	1	1200	555	1955

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 61.47_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 67.01_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 61.47_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 28

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	Sqft_above		Sqft_basement	
1	800	✖	770	✖
2	1200	✖	900	✖
3	1710	✖	234	✖
4	2400	✖	600	✖
5	2800	✖	890	✖
6	1640	✖	489	✖
7	1210	✖	820	✖
8	2530	✖	990	✖
9	990	✖	556	✖
10	790	✖	345	✖
11	790	✖	770	✖
12	1470	✖	330	✖
13	1650	✖	699	✖
14	1980	✖	567	✖
15	1310	✖	1030	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	365000	4	2.5	1570	9600	1	800	770	1972
2	234000	3	1.5	1200	11935	1	1200	900	1968
3	211000	4	2	1710	8288	1.5	1710	234	1970
4	680000	6	3.5	3000	6320	2	2400	600	1969
5	457000	4	3	2800	7845	2	2800	890	2002
6	349000	3	2.25	1640	7261	2	1640	489	1989
7	360000	3	2.75	2030	6840	1	1210	820	1979
8	350000	4	2.5	2530	10155	2	2530	990	1998
9	208000	3	1	990	7420	1	990	556	1960
10	250275	2	1	790	11234	1	790	345	1942
11	435000	2	1	790	11234	1	790	770	1942
12	270000	3	1.75	1800	7763	1	1470	330	1984
13	438000	3	1.75	1650	12940	1	1650	699	1961
14	659000	4	2	1980	23625	2	1980	567	1938
15	645000	4	2.75	2340	6350	1	1310	1030	1974
16	685100	4	2.5	2790	5423	2	2790	344	1999
17	350000	7	3	2800	9569	1	1400	1400	1963
18	375000	2	1	1840	81892	1	1840	300	1955
19	170000	3	1.5	1350	81549	1	1350	789	1966
20	880000	4	2.5	3030	3841	3	3030	960	2005
21	410500	4	2.5	1980	5000	2	1980	999	2008
22	780000	4	2.5	3690	13609	2	3690	1000	1996
23	1225000	5	4.5	3732	4426	2.5	2932	800	2001
24	449000	2	1.75	1510	6852	1	1510	788	2005
25	546000	3	1.75	2000	5000	1	1110	890	1921

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 80.35_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 85.12_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 80.35_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 29

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1130	✖	520	✖
2	1440	✖	999	✖
3	1050	✖	900	✖
4	1130	✖	340	✖
5	4230	✖	456	✖
6	1370	✖	600	✖
7	1540	✖	789	✖
8	2090	✖	900	✖
9	2110	✖	270	✖
10	2156	✖	1000	✖
11	2070	✖	345	✖
12	1270	✖	640	✖
13	1260	✖	660	✖
14	2290	✖	1110	✖
15	1400	✖	444	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	320000	3	1.5	1650	9380	1	1130	520	1978
2	210000	2	1.75	1440	5680	1	1440	999	1978
3	245000	2	1	1050	5900	1	1050	900	1950
4	328000	3	1.75	1470	7650	1	1130	340	1983
5	985000	4	2.25	4230	37769	2	4230	456	1989
6	501000	3	1.75	1970	7972	1	1370	600	1976
7	315000	3	2.25	1540	17424	2	1540	789	1992
8	403000	3	2.75	2090	8354	2	2090	900	2012
9	550000	3	2	2380	17950	2	2110	270	1934
10	347500	4	2.5	2156	3562	2	2156	1000	2012
11	365000	4	2.25	2070	8893	2	2070	345	1986
12	588000	3	1	1910	8167	1	1270	640	1951
13	605000	3	1.75	1920	7400	1	1260	660	1977
14	1300000	3	3.25	3400	5979	2	2290	1110	1937
15	233000	3	2	1400	9177	1	1400	444	1974
16	360000	4	2.5	2210	17715	2	2210	555	1997
17	459000	2	1.75	1170	4887	1	1020	150	1929
18	740000	3	2	2100	72745	1	2100	567	1995
19	305000	3	1	1160	9750	1	1160	890	1967
20	395000	3	1	1500	4000	1	900	600	1925
21	560000	4	2.5	2940	6000	2	2940	890	2000
22	650000	4	1.75	2010	5070	1	1300	710	1963
23	449950	3	2.5	2170	4912	2	2170	234	2003
24	517500	3	1.5	1430	5200	1	1250	180	1940
25	435000	4	1.5	1510	6460	1	1070	440	1970

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 90.35_% of variance in dependent variable is explained by independent variables
- 97.89_ % of variance in dependent variable is explained by independent variable
- 91.08_% of variance in independent variables are explained by dependent variables
- 92.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 91.08_% of variance in independent variables are explained by dependent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 30

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1380	✖	900	✖
2	820	✖	1100	✖
3	1730	✖	450	✖
4	1100	✖	300	✖
5	880	✖	344	✖
6	1060	✖	555	✖
7	1530	✖	780	✖
8	1460	✖	1020	✖
9	1490	✖	890	✖
10	3260	✖	400	✖
11	2340	✖	714	✖
12	1450	✖	346	✖
13	2210	✖	265	✖
14	1130	✖	567	✖
15	1100	✖	450	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	795000	2	1	1380	5000	1.5	1380	900	1905
2	195000	2	1	820	5100	1	820	1100	1953
3	480000	4	2	2180	10575	1	1730	450	1950
4	485000	4	2	1400	3600	1	1100	300	1900
5	130000	2	1	880	9000	1	880	344	1928
6	325000	2	1	1060	5703	1	1060	555	1946
7	440000	3	1.75	1530	7245	1	1530	780	1984
8	550000	4	1.75	2480	14782	1	1460	1020	1958
9	310000	3	1.75	1490	7150	1	1490	890	1967
10	530000	3	2.5	3660	39478	2	3260	400	1989
11	390000	4	2.5	2340	8548	2	2340	714	1993
12	243000	3	1.75	1450	12125	1	1450	346	1981
13	639000	4	2.5	2210	9875	2	2210	265	1990
14	215000	4	1	1130	7400	1	1130	567	1969
15	405000	5	1.75	1550	10500	1	1100	450	1961
16	250000	4	1.75	1710	7140	1	1010	700	1968
17	405500	3	1.5	1010	8108	1	1010	890	1954
18	452000	6	2.25	2660	13579	2	2660	344	1937
19	730000	6	2.25	2660	13579	2	2660	1220	1937
20	267000	3	2	1300	9644	1	1300	580	1994
21	289950	3	2.5	1670	6186	2	1670	840	2002
22	565000	2	1	960	4080	1	960	900	1911
23	1578000	4	3.25	4670	51836	2	4670	899	1988
24	636000	3	2.5	2140	2770	2	1770	370	2007
25	350000	5	3	2520	5500	1	1550	970	2004

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 70.35_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 73.87_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 73.87_% of variance in independent variables are explained by dependent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 31

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1350	✖	520	✖
2	1010	✖	999	✖
3	1120	✖	230	✖
4	1200	✖	890	✖
5	4370	✖	450	✖
6	970	✖	470	✖
7	2580	✖	167	✖
8	3190	✖	267	✖
9	1750	✖	630	✖
10	1420	✖	900	✖
11	2990	✖	800	✖
12	2680	✖	1000	✖
13	2730	✖	455	✖
14	1220	✖	700	✖
15	1500	✖	280	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	497000	3	2.25	1870	9315	1	1350	520	1975
2	280000	2	1	1010	3000	1	1010	999	1925
3	188000	2	1.5	1120	17487	1	1120	230	1924
4	478500	3	1	2090	4755	1	1200	890	1969
5	1145000	4	3.5	4370	18361	2	4370	450	2001
6	222000	3	2	1440	7187	1	970	470	1981
7	799000	4	3	2580	209523	2	2580	167	1984
8	609900	4	2.5	3190	7399	2	3190	267	2006
9	413000	4	3.5	2380	5809	2	1750	630	1995
10	400000	4	1	1420	4875	1.5	1420	900	1930
11	802541	5	2.75	2990	6768	2	2990	800	2014
12	320000	3	2.5	2680	7757	2	2680	1000	1990
13	507000	4	2.5	2730	7649	2	2730	455	2006
14	387500	3	1	1220	8329	1	1220	700	1946
15	396000	3	1	1500	12616	1	1500	280	1967
16	373000	3	1.75	1830	11788	1	1430	400	1958
17	1484900	5	2.5	4570	19252	2	4570	555	1965
18	359900	4	2.25	2470	7698	2	2470	349	1983
19	410000	2	1	800	4342	1	670	130	1927
20	690000	4	2.5	2900	23488	2	2900	890	1992
21	245000	3	2	1410	5760	1	1410	344	1985
22	700000	5	2.5	2830	25958	1	1610	1220	1979
23	268000	4	1.75	1680	9966	1	1100	580	1977
24	225000	3	1.75	1760	26055	1	920	840	1979
25	980000	4	2.25	2240	11034	2	2240	900	1976

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 80.35_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 81.57_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 81.57_% of variance in independent variables are explained by dependent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 32

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1000	✖	900	✖
2	1550	✖	470	✖
3	1120	✖	800	✖
4	1380	✖	770	✖
5	1470	✖	345	✖
6	3180	✖	739	✖
7	2330	✖	1600	✖
8	990	✖	900	✖
9	1970	✖	800	✖
10	1250	✖	1000	✖
11	1140	✖	455	✖
12	1550	✖	700	✖
13	1640	✖	280	✖
14	1255	✖	230	✖
15	1290	✖	540	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	220000	3	1	1000	6020	1	1000	900	1944
2	521000	4	1.75	2020	36400	1	1550	470	1976
3	325000	3	1	1920	6862	1	1120	800	1952
4	597500	3	2	2150	5400	1.5	1380	770	1911
5	337000	3	2.5	1470	3976	2	1470	345	1999
6	1000000	3	2.5	3180	10492	2	3180	739	1991
7	3070000	3	2.5	3930	55867	1	2330	1600	1957
8	495000	3	1	990	6000	1	990	900	1943
9	748000	4	3.5	2770	3330	2	1970	800	2001
10	888000	3	1.5	1250	8710	1	1250	1000	1953
11	269900	3	1.75	1140	22267	1	1140	455	1986
12	496500	4	2.5	2250	14440	1	1550	700	1967
13	475000	3	2.5	1640	5097	2	1640	280	1969
14	312000	3	1.5	1255	1374	3	1255	230	2004
15	323000	3	1	1290	12231	1	1290	540	1976
16	331000	3	1.75	1300	9079	1	1300	370	1986
17	306000	3	1.5	1220	1086	3	1220	480	2007
18	655000	4	2.5	2630	4000	3	2630	560	2002
19	290000	3	1.5	1950	15954	1	1950	120	1959
20	1250000	4	3.25	3160	5000	2	2360	800	1965
21	460000	3	1.75	1550	4708	1	860	690	1949
22	435000	4	2.5	1700	6380	1	850	850	1940
23	664000	4	2.75	2510	11880	1	1630	880	1978
24	865000	4	1.75	1800	4180	2	1800	450	1921
25	385000	3	1.75	1800	18000	1	1200	600	1968

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 90.77_% of variance in dependent variable is explained by independent variables
- 97.89_ % of variance in dependent variable is explained by independent variable
- 95.12_% of variance in independent variables are explained by dependent variables
- 92.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 90.77_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 33

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	820	✖	140	✖
2	1660	✖	799	✖
3	1330	✖	570	✖
4	990	✖	280	✖
5	3250	✖	230	✖
6	960	✖	540	✖
7	960	✖	370	✖
8	1600	✖	480	✖
9	2130	✖	560	✖
10	770	✖	120	✖
11	1070	✖	600	✖
12	1200	✖	900	✖
13	1360	✖	340	✖
14	2050	✖	345	✖
15	2000	✖	1000	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	289500	3	1	960	6400	1	820	140	1944
2	559950	3	2.5	1660	1458	3	1660	799	2014
3	1042500	3	1.75	1900	9375	1	1330	570	1941
4	543500	3	2.25	1270	2790	1	990	280	1970
5	480000	4	2.25	3250	34293	2	3250	280	1983
6	218000	3	1	960	9633	1	960	540	1982
7	565000	2	1.75	1330	6000	1	960	370	1914
8	426700	3	1.75	2080	7700	1	1600	480	1968
9	650000	3	2.5	2690	11575	1	2130	560	1957
10	542000	2	1	890	3060	1	770	120	1910
11	415000	3	1.75	1070	8000	1	1070	600	1977
12	214000	3	1	1200	8316	1	1200	900	1953
13	245000	3	1.5	1360	8910	2	1360	340	1949
14	314500	3	2	2050	13303	1	2050	345	1993
15	1120000	3	1.5	3000	5750	2	2000	1000	1924
16	378000	2	2	1300	3850	1	800	500	1963
17	305000	4	2.5	2250	9091	1	1340	910	1982
18	595000	5	2.75	3740	6750	1	1980	1760	1978
19	729500	3	2.5	1660	1091	3	1530	130	2015
20	430000	3	2.5	2360	6699	2	2360	900	2003
21	240000	3	1.75	1760	6500	1	1150	610	1987
22	545000	3	1.75	1680	6200	1.5	1680	890	1916
23	530000	3	1.75	1320	2500	1	870	450	1918
24	157000	1	1	870	26326	1	870	345	1939
25	425000	4	1	1260	7645	1.5	1260	340	1925

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 62.30_% of variance in dependent variable is explained by independent variables
- 67.89_ % of variance in dependent variable is explained by independent variable
- 65.12_% of variance in independent variables are explained by dependent variables
- 60.99__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 62.30_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 34

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1670	✖	600	✖
2	2420	✖	900	✖
3	1430	✖	340	✖
4	1440	✖	345	✖
5	1450	✖	200	✖
6	2100	✖	1450	✖
7	1110	✖	160	✖
8	940	✖	111	✖
9	1300	✖	370	✖
10	2400	✖	700	✖
11	1100	✖	345	✖
12	1230	✖	640	✖
13	1930	✖	730	✖
14	2540	✖	820	✖
15	1280	✖	450	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	507000	2	1.75	1670	6460	1	1670	600	2004
2	399950	4	2	2420	31465	1	2420	900	1984
3	185000	3	1.75	1430	10816	2	1430	340	1994
4	460000	3	1.75	1440	7070	1	1440	345	1980
5	435500	3	2.25	1450	4789	2	1450	200	1985
6	469000	4	2.75	3550	13938	1	2100	1450	1966
7	531000	2	3	1270	1175	2	1110	160	2000
8	443500	2	1	940	6804	1	940	111	1949
9	223000	3	2	1670	6824	1	1300	370	1990
10	590000	4	3.5	3100	15842	2	2400	700	1996
11	340000	2	2.5	1100	1760	3	1100	345	1997
12	650000	3	2	1870	3388	1	1230	640	1925
13	533600	3	2.5	1930	5080	2	1930	730	1990
14	695000	3	2.75	2540	4694	2	2540	820	2005
15	242000	3	1.75	1280	7524	1.5	1280	450	1988
16	425000	3	2.25	1820	8814	1	1280	540	1967
17	525000	3	2.25	1490	9414	2	1490	470	1981
18	485000	3	2.5	1880	5502	2	1880	900	1988
19	175000	3	1	1250	9775	1	1250	780	1971
20	210000	3	2	1440	7210	1	1440	820	1983
21	260000	3	2.5	1920	7415	2	1920	345	1989
22	577000	3	2.5	2060	5750	1	1330	730	1976
23	200000	3	1.5	1210	10588	1	1210	1000	1958
24	605000	2	1.75	1760	5000	1	940	820	1927
25	325000	3	1	990	6750	1	990	450	1947

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 70.88_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 70.88_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 35

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1220	✖	555	✖
2	1805	✖	890	✖
3	4150	✖	1900	✖
4	2320	✖	1000	✖
5	1384	✖	230	✖
6	2720	✖	450	✖
7	1820	✖	890	✖
8	1530	✖	600	✖
9	2060	✖	900	✖
10	1080	✖	340	✖
11	1510	✖	345	✖
12	1960	✖	200	✖
13	3180	✖	900	✖
14	2110	✖	345	✖
15	1330	✖	730	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	300000	2	1	1220	75794	1	1220	555	1963
2	266000	3	2.5	1805	3402	2	1805	890	2009
3	930000	4	4	6050	84942	2.5	4150	1900	2009
4	530000	4	2.5	2320	5493	2	2320	1000	2004
5	180000	3	1	1384	8960	1	1384	230	1965
6	619000	3	2.5	2720	6439	2	2720	450	2005
7	264000	3	1.5	1820	10608	1	1820	890	1962
8	730000	2	2.25	2130	4920	1.5	1530	600	1941
9	439000	4	2.25	2060	10070	2	2060	900	1975
10	268500	4	1.75	1420	7500	1	1080	340	1981
11	330000	3	1.75	1510	15744	1	1510	345	1985
12	640000	3	2.5	2160	4000	1.5	1960	200	1903
13	900000	3	2.5	3180	12600	2	3180	900	1978
14	370000	3	2.25	2110	13300	2	2110	345	1985
15	685000	3	2	2060	2900	1.5	1330	730	1931
16	505000	3	1.5	1830	217800	1	1010	820	1981
17	265000	3	1.75	1820	32666	1	1820	450	1966
18	290000	3	2.5	1610	2937	2	1610	710	2006
19	296000	4	2.5	1780	3600	2	1780	500	2006
20	360000	4	2.5	1840	9611	2	1840	578	1996
21	240500	3	1.75	1460	10584	1	990	470	1997
22	372500	2	2.5	1400	2958	2	1400	900	2007
23	595000	3	2.5	3370	14402	2	3370	780	2004
24	515000	4	3	3590	8249	2	3590	820	1997
25	1220000	4	2.25	3200	15367	2	3200	345	1962

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 74.60_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 74.60_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 36

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1140	✖	1040	✖
2	3580	✖	800	✖
3	1790	✖	1370	✖
4	750	✖	1470	✖
5	1300	✖	300	✖
6	2990	✖	200	✖
7	2860	✖	1780	✖
8	1400	✖	1010	✖
9	2750	✖	900	✖
10	2050	✖	780	✖
11	1410	✖	820	✖
12	2340	✖	345	✖
13	2870	✖	823	✖
14	4120	✖	450	✖
15	1370	✖	710	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	705000	3	1.75	2180	10221	1	1140	1040	1946
2	1450000	3	3	4380	6320	2	3580	800	1952
3	451000	3	1.75	1790	9813	2	1790	1370	1949
4	261000	2	1	750	8125	1	750	1470	1943
5	396000	3	1.5	1300	8280	1	1300	300	1956
6	725000	4	2.5	2990	29170	2	2990	200	1999
7	1950000	7	3.5	4640	15235	2	2860	1780	1965
8	238000	4	1	1400	7242	1.5	1400	1010	1959
9	750000	6	1.75	2750	9563	2	2750	900	1973
10	331500	4	2.5	2050	10271	2	2050	780	1998
11	460000	4	3	2230	108900	1	1410	820	1978
12	404950	4	2.25	2340	217014	1	2340	345	1982
13	880000	3	2.5	2870	5163	2	2870	823	2014
14	863000	4	2.5	4120	22370	2	4120	450	1997
15	638700	3	1.75	2080	9162	1	1370	710	1960
16	435000	3	2.5	1840	5680	2	1840	500	2008
17	260000	3	1	1180	5350	1.5	1180	578	1959
18	671500	5	2.75	2160	4000	1.5	1530	630	1927
19	610000	3	1	1180	3400	1.5	1180	930	1907
20	405000	2	0.75	1160	15029	1	870	290	1937
21	338900	4	2.5	1830	5612	2	1830	890	2003
22	720000	4	4	3200	7708	2	3200	690	2004
23	442000	5	3	2560	5445	1.5	1760	800	1927
24	650000	5	2.75	2550	5040	1.5	2550	900	1902
25	263500	3	1.75	1330	9917	1	1040	290	1981

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ (1 Mark)

9. p-value slope for yr_built = ✖ (1 Mark)

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 88.39_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 85.12_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 88.39_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 37

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1240	✖	360	✖
2	3040	✖	160	✖
3	6050	✖	720	✖
4	1590	✖	555	✖
5	2500	✖	850	✖
6	1320	✖	460	✖
7	1360	✖	400	✖
8	4740	✖	1000	✖
9	1010	✖	1370	✖
10	1040	✖	1470	✖
11	1610	✖	300	✖
12	1570	✖	200	✖
13	1010	✖	500	✖
14	1360	✖	300	✖
15	1970	✖	234	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	727500	2	2	1240	9119	1	1240	360	1952
2	650000	4	2.5	3040	6587	2	3040	160	2003
3	1600000	6	5	6050	230652	2	6050	720	2001
4	365000	4	1	1590	7085	1.5	1590	555	1944
5	662000	3	1.75	2500	36947	1	2500	850	1984
6	270500	3	2.5	1780	7848	1	1320	460	1978
7	200000	2	1.5	1360	1898	2	1360	400	1978
8	970000	4	3.25	4740	76230	2	4740	1000	1987
9	160000	3	1.5	1010	9600	1	1010	1370	1962
10	204000	3	1	1040	7405	1	1040	1470	1971
11	358000	3	2.25	1610	6655	2	1610	300	1985
12	452000	4	1.75	1570	8268	1	1570	200	1979
13	595000	2	2	1510	4000	1	1010	500	1900
14	249000	3	2	1360	6082	1	1360	300	1994
15	346900	4	2.5	1970	5106	2	1970	234	1999
16	2125000	3	2.5	5403	24069	2	5403	450	1976
17	280000	4	2.5	1560	4350	2	1560	900	2003
18	318888	4	1.75	2320	12000	1	2270	50	1973
19	530000	3	2.25	2120	40518	1	2120	200	1967
20	375000	3	2.25	2120	18500	2	2120	1010	1983
21	910000	4	2.5	3350	29242	2	3350	900	1988
22	782000	4	2.5	2380	9614	2	2380	780	1991
23	259950	2	2	1070	649	2	720	350	2008
24	315000	4	2.75	2300	18360	1	1560	740	1979
25	215000	2	1	1320	8865	1	1320	900	1943

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 86.91_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 85.12_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 86.91_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 38

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1890	✖	100	✖
2	1080	✖	530	✖
3	850	✖	200	✖
4	630	✖	450	✖
5	1780	✖	1130	✖
6	850	✖	850	✖
7	1950	✖	234	✖
8	2990	✖	555	✖
9	900	✖	450	✖
10	4860	✖	890	✖
11	1460	✖	700	✖
12	3860	✖	1000	✖
13	890	✖	670	✖
14	2810	✖	345	✖
15	3400	✖	900	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	520000	4	2.25	1890	3006	2	1890	100	2003
2	460000	3	2.5	1610	2527	2	1080	530	2005
3	438924	3	1.5	1050	4590	1	850	200	1949
4	315000	2	1	630	6550	1	630	450	1918
5	590000	4	2.75	2910	10650	1	1780	1130	1975
6	378500	4	1.75	1700	8640	1	850	850	1955
7	328500	3	2.5	1950	8130	2	1950	234	1990
8	849950	3	3	2990	9773	2	2990	555	1973
9	525000	3	1.5	1350	6000	1	900	450	1950
10	1385000	4	3.25	4860	181319	2.5	4860	890	1993
11	295000	4	2.5	2160	7725	1	1460	700	1978
12	1067000	6	3.5	4860	11793	2	3860	1000	1998
13	207950	2	2	890	5000	1	890	670	1917
14	699900	4	2.75	2810	7302	2	2810	345	2002
15	1280000	5	2.5	3400	9500	2	3400	900	1977
16	452000	4	1.5	1580	7350	1	960	620	1963
17	370000	3	2	1680	5036	1	1680	700	1987
18	232000	3	1.75	1300	11230	1	1300	600	1968
19	669950	5	2.5	2820	14062	2	2380	440	1960
20	397500	3	2	1510	6710	1	1070	440	1972
21	490000	2	2.5	1230	1391	2	870	360	2004
22	725000	4	2	2110	4140	2	1710	400	1925
23	299000	3	2.75	3080	19635	1	1610	1470	1958
24	625000	2	1.5	1490	5750	1.5	1190	300	1900
25	437500	3	2.5	2120	4500	2	2120	200	2000

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 85.30_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 80_% of variance in independent variables are explained by dependent variables
- 82.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 85.30_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 39

Not answered

Marked out of 68

Question1: Spearman's Rank Correlation (Use the Sample data mentioned here)

a. Rank the following variables: (30 Marks: 1 Mark Each)

Sr.No.	R1		R2	
1	1850	✖	567	✖
2	1290	✖	550	✖
3	700	✖	790	✖
4	700	✖	790	✖
5	960	✖	890	✖
6	1160	✖	420	✖
7	1260	✖	763	✖
8	970	✖	970	✖
9	1240	✖	1250	✖
10	1890	✖	630	✖
11	920	✖	840	✖
12	4190	✖	567	✖
13	870	✖	300	✖
14	1600	✖	789	✖
15	1800	✖	960	✖

Spearman's correlation coefficient = ✖ **(1 Mark)**

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data except rank correlation where separate data samples are given.

Sr. No.	price	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	254000	2.5	1850	4597	2	1850	567	2003
2	472800	2.25	1840	8400	1	1290	550	1961
3	263000	1.75	1490	3800	1	700	790	1913
4	510000	1.75	1490	3800	1	700	790	1913
5	179950	1	960	10125	1	960	890	1952
6	525000	2.25	1580	7307	1	1160	420	1973
7	479000	2.5	1260	889	3	1260	763	2008
8	651000	2	1940	6440	1	970	970	1940
9	949000	1.75	2490	7834	1	1240	1250	1958
10	392400	2.5	2520	27951	1.5	1890	630	1942
11	585000	1.75	1760	5356	1	920	840	1950
12	460000	2.5	4190	7504	2	4190	567	2004
13	470000	1	1170	4400	1	870	300	1951
14	245000	2.5	1600	4271	2	1600	789	2003
15	500000	2	2760	27631	2	1800	960	1978
16	455000	2.5	2090	8653	2	2090	999	1989
17	485500	2.5	2320	3974	2	2320	1000	1999
18	1875000	3.25	4110	7920	2	3150	960	1921
19	100000	1	770	17334	1	770	900	1978
20	307000	1.75	1410	16105	1	1410	347	1982
21	475000	3.5	3660	14401	2	2660	1000	1994
22	330000	2.5	2180	7210	1	1190	990	1972
23	130000	1	1110	7520	1	1110	556	1960
24	320000	2.5	2280	7417	2	2280	345	1998
25	735000	2.5	2820	6180	2	2050	770	2013

Question 2: Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 3: Simple Regression (Use original data)

If we take selling price and km driven then

a. Intercept = ✖ (1 Mark)

b. Slope = ✖ (1 Mark)

c. If sqft_living is 3480, then price of house will be = ✖ (1 Mark)

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ (2 Marks)
2. Bathroom ✖ (2 Marks)
3. Sqft_Living ✖ (2 Marks)
4. Sqft_lot ✖ (2 Marks)
5. Sqft_above ✖ (2 Marks)
6. Sqft_basement ✖ (2 Marks)
7. Year_built ✖ (2 Marks)

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ (1 Mark)
2. p-value slope for Bedroom = ✖ (1 Mark)
3. p-value slope for Bathroom = ✖ (1 Mark)
4. p-value slope for Sqft_Living = ✖ (1 Mark)

5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 66.08_% of variance in dependent variable is explained by independent variables
- 67.89_ % of variance in dependent variable is explained by independent variable
- 65.12_% of variance in independent variables are explained by dependent variables
- 62.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 66.08_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 40

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above A		sqft_basement	
1	3180	✖	800	✖
2	860	✖	630	✖
3	1930	✖	700	✖
4	2770	✖	600	✖
5	1670	✖	440	✖
6	1330	✖	440	✖
7	1210	✖	360	✖
8	1030	✖	160	✖
9	820	✖	720	✖
10	4030	✖	555	✖
11	1580	✖	850	✖
12	2600	✖	345	✖
13	4410	✖	900	✖
14	1330	✖	620	✖
15	950	✖	950	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	650000	4	2.5	3180	5438	2	3180	800	2003
2	475000	2	1	1490	3825	1	860	630	1929
3	342000	3	2	1930	11947	1	1930	700	1966
4	1197350	4	2.5	2770	7800	2	2770	600	1999
5	280000	3	1.5	1670	11610	1	1670	440	1963
6	240000	2	1.75	1330	7200	1	1330	440	1993
7	529000	3	1	1210	3328	1.5	1210	360	1924
8	442000	3	1	1190	5100	1	1030	160	1941
9	615000	2	1	1540	6872	1	820	720	1946
10	1485000	4	3.75	4030	10800	2	4030	555	2006
11	585000	4	1.75	2430	7559	1	1580	850	1981
12	480000	4	2.5	2600	7787	2	2600	345	2005
13	1095000	5	3.5	4410	57063	2	4410	900	1990
14	290000	4	1	1330	8184	1.5	1330	620	1949
15	615000	3	1.75	1900	5000	1	950	950	1951
16	300000	2	1	1100	7680	1	1100	900	1950
17	308000	3	1	1980	39150	1.5	1580	400	1932
18	750000	3	2.5	2400	7745	2	2400	1000	2001
19	2250000	4	2.25	4760	8036	2.5	3390	1370	1916
20	890000	4	3.25	3180	194278	2	3180	1470	2005
21	288000	3	1	1190	7560	1	1190	300	1958
22	190000	2	1	630	6000	1	630	200	1943
23	515000	5	3	2530	5105	1	1520	1010	2005
24	235000	4	1.5	1690	11054	1	1690	900	1930
25	275000	2	1	940	5000	1	940	780	1951

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 88.45_% of variance in dependent variable is explained by independent variables
- 87.89_ % of variance in dependent variable is explained by independent variable
- 95.12_% of variance in independent variables are explained by dependent variables
- 92.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 88.45_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 41

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	R1		R2	
1	1500	✖	455	✖
2	2500	✖	899	✖
3	1000	✖	280	✖
4	2420	✖	960	✖
5	2210	✖	700	✖
6	3060	✖	900	✖
7	1330	✖	345	✖
8	3300	✖	678	✖
9	2500	✖	923	✖
10	1570	✖	960	✖
11	1830	✖	567	✖
12	1170	✖	890	✖
13	1680	✖	1250	✖
14	1230	✖	345	✖
15	1620	✖	567	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	505000	2	1500	2500	2	1500	455	2002
2	565000	2.5	2500	7394	2	2500	899	1990
3	585000	1.75	1280	7110	1	1000	280	1955
4	328000	3.25	3380	7500	2	2420	960	1990
5	749950	2.75	2910	18700	1	2210	700	1957
6	826000	2.5	3060	7140	2	3060	900	2006
7	535000	1	1330	40259	1	1330	345	1977
8	925000	2.5	3300	101930	2	3300	678	1991
9	875000	3.5	3250	6000	2	2500	923	2001
10	902000	2.25	2530	9200	1	1570	960	1976
11	423000	1	1830	13900	1	1830	567	1951
12	235000	1	1170	11100	1	1170	890	1968
13	545000	2	2930	14057	1	1680	1250	1980
14	343000	1	1230	7560	1	1230	345	1961
15	334950	2.5	1620	2930	2	1620	567	2002
16	405000	1.75	1760	5355	1	1160	600	1956
17	220000	1	1440	8250	1	1440	300	1959
18	173000	1	1740	8100	1	1050	690	1947
19	655000	1	1270	3600	1.5	1270	344	1906
20	406000	1	880	3000	1	880	555	1924
21	632000	2.5	2560	8320	1	1370	1190	1961
22	357000	2.75	1540	4000	1	1140	400	1990
23	361000	1.75	1150	17585	1	1150	1000	1964
24	277000	1	770	8149	1	770	1500	1948
25	630000	3.75	4610	40202	1	2500	2110	1980

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 45_% of variance in dependent variable is explained by independent variables
- 49.80_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 49.80_ % of variance in dependent variable is explained by independent variable

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 42

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	R1	A	R2	A
1	880	✖	365	✖
2	1330	✖	600	✖
3	1560	✖	1560	✖
4	1830	✖	1250	✖
5	3230	✖	140	✖
6	700	✖	990	✖
7	1010	✖	234	✖
8	1410	✖	1050	✖
9	2230	✖	788	✖
10	1570	✖	899	✖
11	2390	✖	998	✖
12	3020	✖	440	✖
13	1640	✖	760	✖
14	2940	✖	567	✖
15	5990	✖	2020	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	492000	2	1	880	2970	1	880	365	1927
2	425000	4	2.5	1930	14196	1	1330	600	1977
3	1955000	4	2.75	3120	7898	1	1560	1560	1963
4	298900	3	2.5	1830	6162	2	1830	1250	1994
5	699000	4	2.5	3230	40319	2	3230	140	1987
6	310000	2	1	700	3000	1	700	990	1918
7	350000	3	1.75	1010	8580	1	1010	234	1961
8	790000	4	1.75	2460	10061	1	1410	1050	1961
9	305000	4	2.5	2230	5000	2	2230	788	2003
10	315000	3	2.5	1570	2865	2	1570	899	2003
11	267500	5	3.5	2390	6600	2	2390	998	1977
12	1750000	4	3.25	3460	7749	2	3020	440	1950
13	435000	6	3.5	2400	8620	2	1640	760	1987
14	482500	6	4.5	2940	7500	1.5	2940	567	1966
15	5110800	5	5.25	8010	45517	2	5990	2020	1999
16	439995	3	1.75	1570	8400	1	1570	234	1959
17	460000	3	3	1520	7500	1	1180	340	1967
18	194000	3	1	1070	6440	1	1070	567	1971
19	290000	2	0.75	440	8313	1	440	349	1943
20	274700	4	2	2440	9600	1	1220	1220	1963
21	900000	4	2.5	2660	3672	2	1800	860	1910
22	1150000	4	4	4460	103382	2	4460	555	2001
23	605000	3	2.5	2010	3667	2	2010	799	2008
24	359000	3	2.5	1450	3850	2	1450	900	2004
25	850000	3	1.75	1400	9900	1	1400	345	1951

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 93.54_% of variance in dependent variable is explained by independent variables
- 97.89_ % of variance in dependent variable is explained by independent variable
- 95.12_% of variance in independent variables are explained by dependent variables
- 92.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 93.54_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 43

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

R1		R2	
1310	✖	990	✖
1090	✖	920	✖
730	✖	360	✖
1510	✖	810	✖
1040	✖	345	✖
1810	✖	678	✖
1150	✖	890	✖
2360	✖	900	✖
1420	✖	550	✖
2690	✖	234	✖
1040	✖	1040	✖
1040	✖	1040	✖
1370	✖	1620	✖
3410	✖	800	✖
1850	✖	346	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	320000	3	1.75	2300	41900	1	1310	990	1939
2	310000	4	3	2010	7426	1	1090	920	1951
3	356700	2	1	1090	5000	1	730	360	1942
4	775000	4	1.75	2320	5595	1	1510	810	1956
5	325000	3	1	1040	7541	1	1040	345	1951
6	208000	4	2.75	1810	8677	1.5	1810	678	1962
7	825000	2	1	1150	12775	1	1150	890	1908
8	515000	3	2.5	2360	11254	1	2360	900	1990
9	490000	3	1.5	1970	3400	1.5	1420	550	1929
10	675000	4	3	2690	28300	1	2690	234	1954
11	360000	4	2.5	2080	4080	1	1040	1040	1962
12	550000	4	2.5	2080	4080	1	1040	1040	1962
13	225000	3	1.5	1370	9000	1	1370	1620	1962
14	500000	3	2.75	3410	9360	1.5	3410	800	1967
15	385000	4	2	1850	11700	1	1850	346	1969
16	837700	5	2.75	3010	12611	2	3010	789	1994
17	448000	3	2.25	1530	3056	2	1530	200	2003
18	730000	5	1.75	2690	21357	1	1420	1270	1952
19	575000	4	1.5	2810	7140	1	1490	1320	1954
20	430000	3	1.75	1810	7300	1	1240	570	1976
21	435000	3	2.5	1560	3987	2	1560	677	2002
22	241000	3	1	1010	9611	1	1010	900	1973
23	453500	2	1.75	1000	1760	1	600	400	1924
24	285000	2	1.5	990	1380	3	990	500	2001
25	151000	2	1	720	4222	1	720	455	1942

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ **(1 Mark)**

7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 44.1_% of variance in dependent variable is explained by independent variables
- 47.89_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 44.1_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 44

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

R1	A	R2	A
	890		810
	2450		1000
	3440		899
	2380		455
	1460		940
	5670		900
	1140		344
	1270		400
	1280		630
	1560		489
	1500		900
	4475		456
	1640		1200
	1020		1620
	1410		800

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	410000	3	2	1700	4250	1	890	810	1944
2	850000	4	3.25	3450	6500	2	2450	1000	1994
3	1050000	4	3.25	3440	35021	2	3440	899	1983
4	648000	4	2.5	2380	13435	2	2380	455	1992
5	620000	4	2.25	2400	5350	1.5	1460	940	1929
6	935000	4	4.5	5670	84267	2	5670	900	2008
7	379000	3	1	1140	10320	1	1140	344	1963
8	699000	3	1.75	1670	5375	1	1270	400	1952
9	528000	4	2.25	1910	8005	1	1280	630	1960
10	425000	3	1.75	1560	9452	1	1560	489	1974
11	249500	2	1.75	1500	8645	1	1500	900	1963
12	1035480	5	3.25	4475	6642	2	4475	456	2014
13	790000	4	2.75	2840	11900	1	1640	1200	1961
14	399950	3	1	1020	18050	1	1020	1620	1969
15	275000	3	1.5	1410	9000	1	1410	800	1953
16	550000	3	1	1520	2500	1.5	1520	346	1912
17	140000	1	1	730	6890	1	730	789	1926
18	250000	3	1	1190	6250	1	990	200	1954
19	318000	3	2.25	1410	8909	2	1410	356	1988
20	505000	2	1	1240	57000	1	1240	265	1962
21	470500	5	2.5	2500	9248	1	1300	1200	1966
22	360000	3	1	1050	9206	1.5	1050	864	1954
23	1242500	4	2.75	2680	8500	1	1480	1200	1962
24	986000	3	2.5	2380	16080	1	1340	1040	1964
25	515000	2	1	1060	4228	1	860	200	1906

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope: ✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**

6. p-value slope for floor = ✖ (1 Mark)

7. p-value slope for sqft_above = ✖ (1 Mark)

8. p-value slope for sqft_basement = ✖ (1 Mark)

9. p-value slope for yr_built = ✖ (1 Mark)

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 65.50_% of variance in dependent variable is explained by independent variables
- 67.89_ % of variance in dependent variable is explained by independent variable
- 45.12_% of variance in independent variables are explained by dependent variables
- 42.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 65.50_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 45

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr. No.	sqft_above		sqft_basement	
1	1930	✖	899	✖
2	870	✖	478	✖
3	1300	✖	900	✖
4	1490	✖	345	✖
5	1030	✖	350	✖
6	1190	✖	1000	✖
7	2870	✖	1380	✖
8	1920	✖	234	✖
9	2080	✖	578	✖
10	2690	✖	530	✖
11	2060	✖	1620	✖
12	1270	✖	800	✖
13	1480	✖	346	✖
14	2000	✖	789	✖
15	1560	✖	567	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	315000	4	2.5	1930	9643	2	1930	899	1992
2	271310	2	1	870	5340	1.5	870	478	1906
3	315000	4	1	1300	8400	1.5	1300	900	1953
4	245000	3	1.75	1490	6930	1	1490	345	1990
5	543115	2	1	1380	5484	1	1030	350	1947
6	900000	4	2	1190	8190	1.5	1190	1000	1945
7	2450000	4	4.25	4250	6552	2	2870	1380	2008
8	554000	2	2	1920	6045	1	1920	234	2003
9	249900	3	1.75	2080	12522	1	2080	578	1950
10	768000	3	2.5	3220	54160	2	2690	530	1981
11	980000	4	3	3680	5854	1	2060	1620	1967
12	395000	5	2.5	2070	9600	1	1270	800	1962
13	415000	3	1.75	1480	2200	2	1480	346	1995
14	380000	4	2.5	2000	6921	2	2000	789	2003
15	566000	3	2.5	1560	5259	2	1560	567	1999
16	218500	2	1	1600	8961	1	1390	210	1949
17	550000	4	2.5	1940	10500	1	1140	800	1976
18	512031	3	1.75	1540	3000	1	770	770	1920
19	197500	2	1	820	8860	1	820	780	1950
20	470000	3	2.25	1780	8784	1	1230	550	1977
21	590000	3	1.75	1560	242629	1	1560	445	1981
22	465000	4	2.25	2210	8862	1	1270	940	1977
23	499000	2	1.75	2040	114562	1	2040	456	1968
24	515000	3	2.25	1920	11500	1	1920	867	1972
25	716500	3	1.75	1930	5000	1	1230	700	1936

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 73.17_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 73.17_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

Question 46

Not answered

Marked out of 68

Q1. Spearmen's Rank Correlation (Use sample data mentioned here)

(a) Rank the following variable sqft_above and sqft_basement. (30 Marks: 1 Mark Each)

Sr.No.	sqft_above		sqft_basement	
1	1830	✖	566	✖
2	1560	✖	899	✖
3	1230	✖	500	✖
4	1930	✖	1000	✖
5	1422	✖	500	✖
6	1380	✖	1590	✖
7	2430	✖	310	✖
8	2610	✖	788	✖
9	2000	✖	120	✖
10	490	✖	430	✖
11	1050	✖	700	✖
12	1180	✖	840	✖
13	2460	✖	577	✖
14	1780	✖	899	✖
15	1080	✖	456	✖

(b) Calculate spearmen's correlation coefficient between Sqft Above and Sqft basement (Use Original Data) (1mark)

✖

Q2. Simple regression (Use Original Data)

In the below (Original) data, several variables are given where the price is appeared to be affected by other variables. Give the answers of all the below questions based on this set of Original data **except rank correlation where separate data samples are given.**

Sr. No.	price	bedrooms	bathrooms	sqft living	sqft_lot	floors	sqft_above	sqft_basement	yr_built
1	700000	4	2	1830	4590	2	1830	566	1908
2	270000	3	2	1560	8853	1	1560	899	1967
3	258000	4	1.75	1730	8320	1	1230	500	1977
4	530000	3	2.5	1930	7214	2	1930	1000	2005
5	245000	3	2.25	1422	3677	2	1422	500	2012
6	563000	3	2	1380	5000	1.5	1380	1590	1917
7	515000	3	2.5	2430	4203	2	2430	310	2001
8	325000	4	2.5	2610	7091	2	2610	788	1987
9	540000	3	2.25	2000	217800	2	2000	120	1996
10	280300	2	1	920	5000	1	490	430	1949
11	460000	4	1.75	1750	5500	1.5	1050	700	1926
12	475000	3	1.75	2020	8970	1	1180	840	1981
13	900000	4	2.25	2460	44431	1	2460	577	1957
14	538200	4	3	1780	7260	1	1780	899	1964
15	429000	2	1	1080	3600	1	1080	456	1922
16	164808	3	1	1250	5411	1	1250	784	1980
17	520000	4	2.25	1870	7700	2	1870	560	1977
18	475000	3	1.75	1700	8432	1	1230	470	1977
19	283000	4	1.5	1480	47045	1	1480	678	1942
20	321500	4	2.5	1930	6228	2	1930	789	2000
21	840000	4	2.75	3130	21810	2	3130	455	1993
22	287000	3	2.25	1370	6000	2	1370	570	1992
23	375000	3	1.75	960	8106	1	960	360	1962
24	349500	3	1.75	1260	7128	1	1260	344	1969
25	446450	3	1.5	1480	7749	1	1480	555	1960

If we take price and sqft_living, then calculate:

(A) Intercept: ✖

(B) Slope:✖

(C) If sqft_living is 3480, then price of house will be ✖

Q3. Pearson's correlation (Use original data)

Pearson correlation coefficient r between sqft_living & sqft_above = ✖ **(1 Mark)**

Question 4: Multiple Regression (Use original data)

Give the answers of all the below questions of multiple regression based on the original data:

a. Which of the following independent variables are significant? (14 marks)

1. Bedroom ✖ **(2 Marks)**
2. Bathroom ✖ **(2 Marks)**
3. Sqft_Living ✖ **(2 Marks)**
4. Sqft_lot ✖ **(2 Marks)**
5. Sqft_above ✖ **(2 Marks)**
6. Sqft_basement ✖ **(2 Marks)**
7. Year_built ✖ **(2 Marks)**

b. From the selected independent variables answer the following questions: (9 MARKS)

1. p-value intercept = ✖ **(1 Mark)**
2. p-value slope for Bedroom = ✖ **(1 Mark)**
3. p-value slope for Bathroom = ✖ **(1 Mark)**
4. p-value slope for Sqft_Living = ✖ **(1 Mark)**
5. p-value slope for Sqft_Lot = ✖ **(1 Mark)**
6. p-value slope for floor = ✖ **(1 Mark)**
7. p-value slope for sqft_above = ✖ **(1 Mark)**

8. p-value slope for sqft_basement = ✖ **(1 Mark)**

9. p-value slope for yr_built = ✖ **(1 Mark)**

Question 5: Answer the following question: From the value of R^2: (5 Marks) (Use original data)

- 73.07_% of variance in dependent variable is explained by independent variables
- 77.89_ % of variance in dependent variable is explained by independent variable
- 75.12_% of variance in independent variables are explained by dependent variables
- 72.45__ % of variance in independent variables are explained by dependent variable

Mark 0 out of 5

The correct answer is: 73.07_% of variance in dependent variable is explained by independent variables

Question 6: Based on the above information, (Model 1&2) which of the following is true? (5 marks)

	Model 1	Model 2
Variables Used	X1,X2,X3,Y1	X1,X2,Y1
R-Square	0.5923	0.5612
Adjusted R Squared	0.4431	0.3512

- Comparing the values of R Squared & adjusted R Square, Model 1 is better.
- Comparing R squared between Model 1 & 2, the R squared predicts Model 2 is better as it carries greater explanatory power.
- Comparing the adjusted R Square between Model 1 & 2 , it predicts that the input variable X3 does not contribute to explaining output variable Y1.
- Comparing values of R squared & adjusted R Square, Model 2 is better.

Mark 0 out of 5

The correct answer is: Comparing the values of R Squared & adjusted R Square, Model 1 is better.

PREVIOUS ACTIVITY
STA 101_MID SEM EXAM_ 2021 (hidden)