KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
MATHEMATICS DEPARTMENT

CODE 000

STAT 212 BUSINESS STATISTICS II
Semester 233, Final Exam
August 15, 2023

CODE 000

Time allowed 120 minutes

Name: _____ ID #:_____

Important Instructions:

Check that this exam has 20 questions

- All types of mobile phones or smart watches are NOT allowed during the examination.
- Use HB 2.5 pencils only.
- Use a good eraser. DO NOT use the erasers attached to the pencil.
- Write your name, ID number and Section number on the examination paper and in the upper left corner of the answer sheet.
- When bubbling your ID number and Section number, be sure that the bubbles match with the numbers that you write.
- The Test Code Number is already bubbled in your answer sheet. Make sure that it is the same as that printed on your question paper.
- When bubbling, make sure that the bubbled space is fully covered.
- When erasing a bubble, make sure that you do not leave any trace of penciling
- Formula sheet will be provided to you in exam. You are not allowed to bring, with you, formula sheet or any other printed/written paper.

Ex 14.5, 14.13, 14.27

How does horsepower and weight affect the mileage of family sedans? Data from a sample of twenty 2010 family sedans were collected and the results of the regression model and ANOVA table shown in the following output

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	27.8053	13.9026	7.5886	0.0044
Residual	17	31.1447	1.8320		
Total	19	58.9500			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	41.5895	6.3089	6.5922	0.0000	28.2788	54.9001
Horsepower	-0.0222	0.0105	-2.1257	0.0485	-0.0443	-0.0002
Weight	-0.0041	0.0020	-2.0074	0.0609	-0.0084	0.0002

Based on the above Answer the next **six** questions

1. For a given weight, each increase of one unit in horsepower is estimated to result in a mean decrease in MPG of _____.

   A. 0.0222.
   B. – 0.0222.
   C. 0.0041.
   D. – 0.0041.
   E. No practical meaning here because it would have meant the estimated mean gasoline mileage when a car has 0 horsepower and 0 weight.

2. Predict the miles per gallon for cars that have 190 horsepower and weigh 3,500 pounds is:

   A. 23.0215 MPG
   B. 41.0242 MPG
   C. 37.2280 MPG
   D. 39.7327 MPG
   E. 31.5222 MPG

3. The percentage of the variation in MPG can be explained by variation in horsepower and variation in weight, taken in account the sample size and the independent variables is

   A. 40.95%.
   B. 41.17%.
   C. 27.81%.
   D. 58.95%.
   E. 31.14%.

4. A 95% confidence interval estimate of the population slope between mileage and horsepower:

   A. Between – 0.0443 and – 0.0002
   B. Between – 0.0084 and 0.0002
   C. Between 28.2788 and 54.9001
   D. Between – 0.0405 and – 0.00393
   E. Between – 0.00758 and – 0.00062

5. At 5% level of significance, which of the following statements is NOT true?

   A. Only the Weight is significant
   B. Only the Horsepower is significant
   C. Only the Weight is not significant
   D. The overall model is significant
   E. There is no meaning of the MPG when the horsepower is zero and weight is zero

6. Given that $SSR_{(horsepower)}$ =20.4225 and $SSR_{(weight)}$ = 19.5268, the proportion of variation in the MPG is explained by the horsepower while controlling for (holding constant) the weight.

   A. 20.99%
   B. 19.16%
   C. 47.17%
   D. 40.95%
   E. 8.278%

Ex 14.41

7. The marketing manager of a large supermarket chain faced the business problem of determining the effect on the sales of pet food of shelf space (X_1) and whether the location of the product (X_2) was placed at the front $(X_2 = 1)$ or back $(X_2 = 0)$ of the aisle. Data are collected from a random sample of 12 equal-sized stores. believes that the appropriate model is

$$\hat{y} = 130 + 7.4\,X_1 + 45X_2$$

The predicted sales for the product when the shelf space is 8.5 feet, and it is placed at the front is:

   A. 237.9
   B. 192.9
   C. 182.9
   D. 229.9
   E. 259.9

You constructed simple linear regression models to investigate the relationship between monthly sales for a chain of sporting goods stores and five independent variables.
Based on the following MINITAB output

Best Subsets Regression: y versus x1, x2, x3, x4, x5

```
Response is y

                 R-Sq    R-Sq  Mallows                 x x x x x
Vars  R-Sq     (adj)   (pred)      Cp         S        1 2 3 4 5
   1  24.1     21.9    17.2      -1.8    802004           X
   1  14.7     12.3     4.0       2.1    849860         X
   2  24.3     20.0     8.7       0.0    811809   X     X
   2  24.2     19.9    14.0       0.1    812572         X X
   3  24.4     17.8     5.4       2.0    823193   X     X X
   3  24.4     17.7     0.0       2.0    823626   X X   X
   4  24.4     15.3     0.0       4.0    835508   X   X X X
   4  24.4     15.3     0.0       4.0    835563   X X   X X
   5  24.4     12.6     0.0       6.0    848433   X X X X X
```

Correlation: y, x1, x2, x3, x4, x5

```
          y      x1      x2      x3      x4
x1    0.041
x2    0.112  -0.203
x3    0.384   0.387   0.032
x4    0.490   0.193   0.217   0.804
x5    0.322   0.265   0.381   0.682   0.712
```

Based on the above Answer the next **three** questions

8. If only one variable were be brought into the model, which variable should be if the goal is to explain the highest possible percentage of variation in the dependent variable?

   A. X_4 has the highest correlation with monthly sales.
   B. X_1 has the highest correlation with monthly sales.
   C. X_2 has the highest correlation with monthly sales.
   D. X_5 has the highest correlation with monthly sales.
   E. X_3 has the highest correlation with monthly sales.

9. Using R^2_{adj}, the model should be taken into consideration contains

   A. 1 variable .
   B. 4 variables.
   C. 2 variables.
   D. 3 variables.
   E. 5 variables.

10. If you are going to fit a regression model using the backward elimination selection method, what is the first predictor to be removed

   A. X_1
   B. X_2
   C. X_4
   D. X_5
   E. X_3

11. Consider a nine-year moving average used to smooth a time series that was first recorded in 2002. How many years of values in the series are lost when computing all the nine-year moving averages?

A. A total of 8 years.
B. A total of 4 years.
C. A total of 6 years.
D. A total of 5 years.
E. A total of 9 years.

12. You are using exponential smoothing on an annual time series concerning total revenues (in millions of dollars). You decide to use a smoothing coefficient of $w = 0.2$, and the exponentially smoothed value for 2010 is $E_{2010} = 0.2(12.1) + 0.8(9.4)$. the smoothed value of this series in 2011 if the value of the series in that year is $11.5 million is

A. 10.252
B. 11.98
C. 10.392
D. 9.96
E. 9.82

13. If the coefficient of determination between two independent variables is 75%, then the VIF equal to

A. 4
B. 5
C. 3
D. 2
E. 1

14. The linear trend forecasting equation for an annual time series containing 42 values (from 1969 to 2010) on net sales (in billions of dollars) is $\hat{Y}_i = 1.2 + 0.5X_i$. The projected trend forecast two years after the last value is

A. 22.7 million dollars
B. 23.2 million dollars
C. 22.2 million dollars
D. 21.7 million dollars
E. 21.2 million dollars

Betts Electronics purchases three replacement parts for robotic machines used in its manufacturing process. Information on the price of the replacement parts and the quantity purchased is given below:

	PRICE		QUANTITY	
PART	2000	2016	2000	2016
A	$ 2.00	$ 5.50	120	220
B	$ 2.50	$ 6.00	90	110
C	$ 4.00	$ 10.00	80	95

Using the above information to answer the following **three** questions:

15. What is the simple price index for part B in 2016, using 2000 as a base year:

A. 240%
B. 86.54%
C. 120%
D. 41.66%
E. 48.96%

$$I_p = \frac{P_{2016}}{P_{2000}} * 100\% = \frac{6}{2.5} * 100\% = 240\%$$

16. The aggregate price index for 2016, using 2000 as a base year is equal to:

A. 252.94%
B. 39.53%
C. 86.45%
D. 210.96%
E. 198.78%

$$I_{ag} = \frac{\sum P_{2016}}{\sum P_{2000}} * 100 = \frac{P_{A,16} + P_{B,16} + P_{C,16}}{P_{A,00} + P_{B,00} + P_{C,00}} = \frac{5.5 + 6 + 10}{2 + 2.5 + 4} * 100 = 252.94\%$$

17. The Laspeyres price index for 2016 using 2000 as the base period is equal to:

A. 254.78%
B. 257.53%
C. 38.83%
D. 39.25%
E. 270.45%

$$I_L = \frac{\sum P_i^t Q_i^0}{\sum P_i^0 Q_i^0} * 100 = \frac{P_{A,16}Q_{A,00} + P_{B,16}Q_{B,00} + P_{C,16}Q_{C,00}}{P_{A,00}Q_{A,00} + P_{B,00}Q_{B,00} + P_{C,00}Q_{C,00}} = \frac{5.5 * 120 + 6 * 90 + 10 * 80}{2 * 120 + 2.5 * 90 + 4 * 80} * 100 = 254.78\%$$

18. In forecasting a quarterly time series over the five year period from the first quarter of 2006 through the fourth quarter of 2010, the exponential trend forecasting equation is given by

$$\log \hat{y} = 3 + 0.1\, X_i - 0.25\, Q1 + 0.2\, Q2 + 0.15\, Q3$$

where quarter zero is the first quarter of 2006. Take the antilog of the appropriate coefficient from this equation, the estimated quarterly compound growth rate in the price is:

A. 25.89%
B. 58.49%
C. 10.51%
D. 41.25%
E. 43.77%

19. A third-order autoregressive model is fitted to an annual time series with 17 years (from 2000 to 2016) and has the following autoregressive equation

$$\hat{Y}_t = 4.5 + 1.8\, Y_{t-1} + 0.8\, Y_{t-2} + 0.24\, Y_{t-3}$$

The three most recent values are $Y_{2014} = 23$, $Y_{2015} = 28$ & $Y_{2016} = 34$, the forecast value for 2018 equal to

A. 206.936
B. 93.62
C. 200.936
D. 205.736
E. 410.9048

20. : Which of the following would be an appropriate alternative hypothesis?

A. The population proportion is less than 60 percent.
B. The sample proportion is less than 60 percent.
C. The population proportion is no less than 60 percent.
D. The sample proportion is no less than 60 percent.
E. The population proportion is no more than 60percent.