Module 6 Assignment

Name: Obubelebara George

PART 1.

1. Seasonal Patterns in Juvenile Crime.

(Use Q1 Data Tab) Early criminological research has found that the frequency of crime is not consistent throughout the year and that crime often peaks in the summer months. However, there is continued debate over which specific crime types experience such trends. In the context of the City of Normalton in Australia, the researcher wants to investigate whether the total number of juvenile criminal arrests varies across seasons.

The dataset includes three variables; the number of total juvenile criminal arrests by month from January 1, 1990 through December 31, 1998 (*Crime*), month of year (*Month*) and year (*Year*).
In Australia, he seasons are defined by grouping the calendar months in the following way:

- Spring - the three transition months September, October and November.
- Summer - the three hottest months December, January and February.
- Autumn - the transition months March, April and May.
- Winter - the three coldest months June, July and August.

Using analysis of variance (ANOVA), determine whether any of these differences in the number of juvenile criminal arrests by season are statistically significant.

- Hypothesis:
- Test result table: (maybe in Excel) - Result:

ANSWER

Hypothesis:

Null Hypothesis (H_0): No difference between any of the group means exists in the population.

Alternative Hypothesis (H_1): At least one difference between group means exists in the population.

Test Result:

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
CRIME - SPRING	27	14792	547.8518519	14144.59259
CRIME - SUMMER	27	12369	458.1111111	8923.102564
CRIME - AUTUMN	27	16451	609.2962963	14019.37037
CRIME - WINTER	27	19550	724.0740741	9124.148148

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1010139.444	3	336713.1481	29.1455793	0.00	2.691979
Within Groups	1201491.556	104	11552.80342			
Total	2211631	107				

Result Interpretation

F vs. Critical Value: We can reject the null hypothesis because the F-value (29.146) is greater than the critical value (2.692).

p-value: We can reject the null hypothesis because the p-value (0.00) is exceedingly small, even less than zero, and less than the α = (0.05) significance level. $P=0.00<0.05$

The analysis indicates a statistically significant difference in the means of juvenile criminal arrests across seasons in the City of Normalton, Australia. Further analysis or posthoc tests are needed to identify which specific seasons differ significantly from each other in terms of juvenile criminal arrest rates, and I tried to do that in the Excel sheet.

2. Unemployment Rate by Population

(Use Q2 Data tab). This dataset is developed from secondary data sources to provide a profile of cities across the US. The 98 selected cities include many of the country's largest, as well as those from different states and a few located in rural settings. The data include unemployment rate of each city (*Unempl*). By population, the researcher ranked the cities and created a new variable named "*Pop category*."

- Pop category =1: Top 1-10 populated cities
- Pop category =2: Top 11-30 populated cities • Pop category =3: Top 31-50 populated cities • Pop category =4: Top 51-75 populated cities
- Pop category =5: Top 76-98 populated cities

The researcher wants to know whether the unemployment rate varies across these categories of top populated cities. Using analysis of variance (ANOVA), determine whether any of these differences are statistically significant.

- Hypothesis:
- Test result table: (maybe in Excel) - Result:

ANSWER

Hypothesis:

Null Hypothesis (H_0): There are no differences between any of the group means in the unemployment rate across categories of top-populated cities.

Alternative Hypothesis (H_1): At least one difference between group means exists in the unemployment rate across categories of top-populated cities.

Test Result

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Unempl - C1	10	100.8	10.08	16.06177778
Unempl - C2	20	175.1	8.755	3.813131579
Unempl - C3	20	162	8.1	5.384210526
Unempl - C4	22	202.1	9.186363636	13.14504329
Unempl - C5	18	137.1	7.616666667	4.039117647

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	52.24681313	4	13.06170328	1.67201407	0.163985905	2.47901547
Within Groups	664.0164091	85	7.811957754			
Total	716.2632222	89				

Result Interpretation:

F vs. Critical Value: We cannot reject the null hypothesis because the F-value (1.672) is less than the critical value (2.479).

p-value: The p-value (0.164) is greater than the significance level of 0.05. Hence, we cannot reject the null hypothesis.

Based on the analysis, we do not have enough evidence to conclude that there is a significant difference between the means of the groups in the unemployment rate across categories of toppopulated cities. Therefore, we fail to reject the null hypothesis.

3. Satisfaction with Working Condition by Years Employed

(Use Q3 Data tab) The Employee Attitudes dataset includes data from 977 employees of Seminole County Government in Florida. This general employee survey includes various aspects of the workplace. Among them, the researcher wants to investigate whether "working conditions" varies across years employed.

- *CONDITION*: You must create a new variable ("conditions") out of the following three variables. You can use "=average(B2, C2, D2)" function. You should delete the errors.
    - PROBLEMS: Problems are discussed openly, candidly, and constructively. (1=Strong disagree; 5=Strongly agree) ○ FAIRSHAR: In general, everyone seems to carry his or her fair share of the workload. (1=Strong disagree; 5=Strongly agree)
    - FREEDOM: I have a lot of freedom to decide how to do my work. (1=Strong disagree; 5=Strongly agree)
- *YEARSERV*:
    - 1=Less then 1 year ○ 2=1+ to 3 years ○ 3=3+ to 5 years ○ 4=5+ to 10 years
    - 5= 10+ years

Remember! First, you need to create the "conditions" variable. Then use analysis of variance (ANOVA) to determine whether "differences" has different averages across "yearserv".

- Hypothesis:
- Test result table: (maybe in Excel) - Result:

Hypothesis:

Null Hypothesis (H_0): No difference between any of the group means exists in the population.

Alternative Hypothesis (H_1): At least one difference between group means exists in the population.

ANSWER

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
CONDITION-YS1	85	302.5	3.558823529	0.544117647
CONDITION-YS2	126	409	3.246031746	1.143428571
CONDITION-YS3	104	343.1666667	3.299679487	0.733753942
CONDITION-YS4	173	552.1666667	3.191714836	0.775208735
CONDITION-YS5	297	945.1666667	3.182379349	0.708273299

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	10.31032	4	2.577579843	3.311142663	0.010564816	2.383347615
Within Groups	607.1959	780	0.778456293			
Total	617.5062	784				

Result Interpretation:

F vs. Critical Value: We can reject the null hypothesis because the F-value (3.311) is greater than the critical value (2.383).

p-value: The p-value (0.011) is less than $\alpha = 0.05$ significance level, Thus, we reject the null hypothesis.

This indicates at least one significant difference between group means in working conditions across years employed. The ANOVA analysis suggests that the length of time employed (yearserv) significantly impacts employees' perceptions of working conditions. In the Excel sheet, I tried to analyze post-hoc tests further to determine which groups differ from each other.

PART 2. Performance Report Project –Citizen Satisfaction Data

Using your variables in your research proposal,
1) Develop ANOVA hypotheses (Mean difference among 3 or more groups)

2) Run ANOVA analysis
3) Report the result

Hypothesis:

Null Hypothesis (H_0): No difference between any of the group means exists in the population.

Alternative Hypothesis (H_1): At least one difference between group means exists in the population.

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
SR – WHITE	1016	4216	4.149606299	0.649517086
SR – BLACK	53	220	4.150943396	0.822931785
SR – ASIAN	244	989	4.053278689	0.667931593
SR – HISPANIC	76	316	4.157894737	0.668070175

Suggested Table for your final paper:

Table 2. ANOVA Results						
Source of Variation	**SS**	**Df**	**MS**	**F**	**P-value**	**F crit**
Between Groups	1.895035645	3	0.631678548	0.956706764	0.412398079	2.611327465
Within Groups	914.4649356	1385	0.660263491			
Total	916.3599712	1388				

Result Interpretation:

F vs. Critical Value: The F-value of 0.957 is less than the critical value (2.611). Thus, we cannot reject the null hypothesis.

ANSWER

p-value: The p-value of 0.412 is greater than the significance level of 0.05. Thus, we cannot reject the null hypothesis.

Based on the ANOVA analysis conducted on the variable "Snow Removal on Major City Streets," This suggests that there is not enough evidence to conclude that there is a significant difference or disparity between the groups' means. Based on the analysis, no statistically significant difference was observed between the means of the groups. Therefore, we do not have enough evidence to reject the null hypothesis.