Module 3 Task

Name: Obubelebara George

PART 1.

1. This dataset (**Watershed** dataset) was developed from rating variables, explanations, descriptions, and specific data incorporated within the Environmental Protection Agency's (EPA) Index of Watershed Indicators.

Open the Watershed dataset, and

a. **Make a pivot table from variables in that data set.**
    a. Consider region (Region) and the quality of watershed (Quality) as the two categorical variables and
    b. show the mean level of conventional pollutants (Conpolut) in each data cell

Variable: Quality (quality of watersheds)

- 1= better quality, low vulnerability
- 2= better quality, high vulnerability
- 3= moderate quality, low vulnerability
- 4= moderate quality, high vulnerability
- 5= serious quality, low vulnerability
- 6= serious quality, high vulnerability

Answer

In completing the Module 2 assignment, I opted to fill in missing values using the average (mean) of the available dataset. This approach was chosen to ensure statistical integrity and to provide a comprehensive perspective for subsequent analyses. By incorporating the average of the available data, the dataset was effectively cleaned, potentially impacting the outcomes of future assignments, such as Module 3. This method diverges from the simple omission of missing values and aims to maintain consistency and coherence in the dataset, potentially influencing the conclusions drawn in subsequent assignments.

Average of Conpolut (Ambient water quality-conventional pollutants)	Column Labels						
Row Labels	1	2	3	4	5	6	Grand Total
1	7.52	8.90	13.50	16.56	9.79	17.42	14.05
2	15.79		19.21	26.70	18.61	26.30	18.59
3	11.68		13.78	19.69	24.64	35.56	20.79
4	6.72		14.85		16.87		15.19

5	10.97	14.55	14.55		17.95	54.28	20.34
Grand Total	**12.34**	**12.67**	**14.76**	**18.34**	**21.55**	**27.28**	**18.34**

Interpretation of the data above:

The average values generally increase from left to right within each row. The Conpolut values tend to increase with higher levels of vulnerability, regardless of the quality level. Additionally, there seems to be a trend where higher quality watersheds (categories 1 and 2) have lower Conpolut averages compared to moderate or serious quality watersheds (categories 3, 4, 5, and 6). The Grand Total row provides the overall average Conpolut value across all quality and vulnerability categories, which is 18.34.

2. Workbook Chapter 8, Data-based exercise #4 (p.48)

The **Employee Attitudes** dataset includes data from 977 employees of Seminole County Government, FL. This survey includes employee assessments of working conditions, career development, benefits and compensation, supervisory management, customer relations, job skills and training, and satisfaction with HR services. Among the dataset, you will use the following variables for this assignment.

- **Dept (What is your department?)**:1=Administrative Services, 2=Community Services, 3=County Attorney's Office, 4=County Manager/BCC, 5=Environmental Services, 6=Fiscal Services, 7=Human Services, 8=Information Technologies, 9=Judicial, 10=Library & Leisure Services, 11=Planning & Development, 12=Public Safety, 13=Public Works, 14=Tourism
- **Gender (What is your gender?)**: 1= Female, 0=Male
- **Race (What is your race?)**: 1=Caucasian, 2= African-American; 3=Asian or Pacific, 4=Native American, 5=Other
- **Bestqual** (Please evaluate the following statement by checking the appropriate answer: The people that get promoted are among the best qualified for the job): 1= strongly disagree, 2=disagree, 3=don't know 4=agree, 5=strongly agree
- **Satisjob** (Please evaluate the following statement by checking the appropriate answer: Overall, I am satisfied with my job at Seminole County): 1= strongly disagree, 2=disagree, 3=don't know 4=agree, 5=strongly agree

2-a. A manager wishes to examine the relationship between race (Race) and perceptions that the people who get promoted are among the best qualified (Bestqual) in the **public works departmen**t. However, very few employees are minorities, and a manager is concerned that separate analysis for each minority group might reveal their identity. Therefore, the manager wants to *compare Caucasian employees against non-Caucasian employees*. Recode the variable Race in this manner and Create a contingency table (Bestqual by Caucasian vs. Non-caucasian). Table should include ***count*** and ***column percentage*** information .

*** TIP: Be sure to analyze only employees in the public works department (dept = 13 under the Filters)*

*** TIP: Here, you should EXCLUDE Blank for both Caucasian and Bestqual*

Answer

	Column Labels											
	1		2		3		4		5		Total Count of race	Total %
Row Labels	Count of race	%	Count of race	%	Count of race	%	Count of race	%	Count of race	%		
Caucasian	24	75.00%	37	72.55%	27	79.41%	35	74.47%	10	76.92%	133	75.14%
Non-Caucasian	8	25.00%	14	27.45%	7	20.59%	12	25.53%	3	23.08%	44	24.86%
Grand Total	**32**	**100.00%**	**51**	**100.00%**	**34**	**100.00%**	**47**	**100.00%**	**13**	**100.00%**	**177**	**100.00%**

Interpretation of the data above:

Across all five columns, the majority of individuals identify as Caucasian, with percentages ranging from 72.55% to 79.41%. On the contrary, the percentage of non-caucasian individuals ranges from 20.59% to 27.45% across the five columns. The data illustrates the racial distribution within a population, indicating that Caucasians comprise the majority with 75.14% (133 individuals), while non-Caucasians account for 24.86% (44 individuals). The contingency table comparing perceptions of promotion fairness between Caucasian and non-Caucasian employees in the public works department reveals potential disparities in their perceptions of whether those promoted are among the best qualified.

2-b. First, make a pivot table from variables in the dataset you create for Q2-a. Bestqual and Caucasian variables (which was used for 2-a) as the two categorical variables, and show the **mean** (average) level of employee satisfaction (Satisjob). Second, **interpret** the table (think about the race-related difference in employee satisfaction, the relationship between Bestqual and Satisjob).

*** TIP: This is for all departments – you may choose all departments under Filters or no need to set up Filters*

Average of satisjob	Column Labels						
Row Labels	**1**	**2**	**3**	**4**	**5**	**(blank)**	**Grand Total**
Caucasian	3.13	3.79	3.99	4.14	4.65	3.75	3.89
Non-Caucasian	3.30	3.75	3.88	4.29	4.75	3.50	3.87
Grand Total	**3.19**	**3.78**	**3.96**	**4.17**	**4.68**	**3.68**	**3.88**

In the process of formulating the Pivot table, it became apparent that there were instances of missing data in the Bestqual column, as evidenced by the blanks in the table provided. Such occurrences suggest that respondents might have opted not to respond to this specific question,

resulting in blank entries. It's noteworthy that in the calculation, I chose to include these blank values. However, for the Pivot table presented below, I opted to exclude the blank entries. This deliberate omission of blanks led to a slight disparity between the two Pivot tables. I also choose to interpret the pivot table below.

Average of satisjob	Column Labels					
Row Labels	1	2	3	4	5	Grand Total
Caucasian	3.13	3.79	3.99	4.14	4.65	3.89
Non-Caucasian	3.30	3.75	3.88	4.29	4.75	3.88
Grand Total	**3.19**	**3.78**	**3.96**	**4.17**	**4.68**	**3.89**

Race-Related Difference in Employee Satisfaction

Caucasian Employees

For those who strongly disagree with the statement that people who get promoted are among the best qualified (Bestqual = 1), the average satisfaction level is 3.13. As the perception of promotion fairness improves (Bestqual increases), the average satisfaction level tends to increase steadily, reaching the highest average satisfaction level of 4.65 for those who strongly agree with the statement (Bestqual = 5). The overall average satisfaction level for Caucasian employees is 3.89.

Non-Caucasian Employees:

Non-Caucasian employees who strongly disagree with the statement about promotion fairness (Bestqual = 1) have a slightly higher average satisfaction level compared to their Caucasian counterparts, with an average of 3.30. Similar to Caucasian employees, as the perception of promotion fairness improves (Bestqual increases), the average satisfaction level tends to increase, reaching a maximum of 4.75 for those who strongly agree with the statement (Bestqual = 5). The overall average satisfaction level for Non-Caucasian employees is 3.88.

The analysis reveals a consistent trend of increased satisfaction with improving perceptions of promotion fairness, irrespective of racial background. Interestingly, Non-Caucasian employees exhibit slightly higher average satisfaction levels across all levels of Bestqual compared to their Caucasian counterparts. However, the disparities in average satisfaction levels between the two racial groups are marginal. These findings underscore the significant influence of perceptions of promotion fairness on employee satisfaction, regardless of racial identity. Although there are

discernible differences in satisfaction levels between Caucasian and Non-Caucasian employees, they do not appear significant enough to warrant considerable concern.

2-c. The county manager wanted to see how satisfied the employees are with their job. In particular, he wanted to know if there is any gender difference in the level of satisfaction in each department. First, make a pivot table, using Dept and Gender as the two categorical variables, and show the **mean** (average) level of employee satisfaction (Satisjob). **Interpret** the table.

Average of satisjob Column Labels			
Row Labels	**0**	**1**	**Grand Total**
1	3.38	3.94	3.58
2	3.18	3.44	3.33
3	4.50	4.50	4.50
4	4.00	4.57	4.50
5	3.68	3.82	3.72
6	4.25	4.44	4.38
7	4.00	3.60	3.71
8	4.36	4.40	4.37
9	4.00	4.27	4.25
10	3.63	4.02	3.85
11	4.02	4.02	4.02
12	3.86	3.84	3.85
13	4.03	4.19	4.05
14	4.50		4.50
(blank)	3.76	3.89	3.84
Grand Total	**3.86**	**3.98**	**3.90**

Here's a brief interpretation of the pivot table:

Each row corresponds to a specific department, identified by a numerical label and each column represents a gender category (0 for male and 1 for female), with the Grand Total column showing the overall average across all genders. The average satisfaction level for employees in department 1 is 3.58. This value is further broken down into two columns representing the mean satisfaction level for each gender category as shown in the first row. It also points out that in row 1 for Gender 0 (male) the mean satisfaction level is 3.38, while for Gender 1 (female), the mean

satisfaction level is 3.94. The Grand Total column and row provide the overall average satisfaction level across all departments and genders, which is 3.90.

The analysis of the pivot table reveals that satisfaction levels differ among departments and genders. It's evident that some departments have higher average satisfaction levels than others. Interestingly, there are slight variations in satisfaction levels between genders within certain departments. For instance, in department 1, the male gender tends to have a slightly higher average satisfaction level compared to the female gender. Additionally, across all departments, we can observe fluctuations in satisfaction levels between genders. It's worth noting that the dataset contained some blank entries in the department column, which are appropriately indicated in the table. Overall, the table offers valuable insights into how employee satisfaction levels vary across departments and genders, laying the groundwork for further exploration of potential gender differences in satisfaction within each department.

In conclusion, the average level of employee satisfaction across all departments and genders being 3.90 suggests that, on average, employees in this organization tend to report a moderate level of satisfaction with their job. This value serves as a benchmark for assessing satisfaction levels within specific departments and genders to determine if there are any notable deviations from the overall average.