**Background Notes:**

Fifty-five Canadians were randomly selected from the Canadian Community Health Survey. Below are five questions that were asked and the responses to each question.

1. *What is your gender?*

*Male*

*Female*

1. *How would you rate your health?*

*Poor*

*Fair*

*Good*

*Very Good*

*Excellent*

1. *What do you think your body mass index (BMI) is?*
2. *What is your Weight (kilograms)?*
3. *How frequent do you perform any type of physical activity?*

*Regular*

*Occasional*

*Infrequent*

**Respondent Data:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID Number | Gender | Health Rating | BMI | Weight | Physical Activity |
| 1 | F | Fair | 24.49 | 60.75 | Regular |
| 2 | M | Excellent | 17.60 | 43.65 | Regular |
| 3 | M | Good | 21.60 | 74.25 | Regular |
| 4 | M | Fair | 20.18 | 75.15 | Missing |
| 5 | F | Poor | 25.40 | 63.00 | Regular |
| 6 | F | Fair | 22.52 | 65.25 | Regular |
| 7 | M | Very Good | 27.49 | 94.50 | Infrequent |
| 8 | M | Poor | 20.76 | 67.50 | Regular |
| 9 | F | Good | 37.62 | 96.30 | Regular |
| 10 | M | Fair | 26.94 | 82.80 | Regular |
| 11 | M | Poor | 24.87 | 85.50 | Occasional |
| 12 | M | Good | 30.29 | 87.75 | Regular |
| 13 | F | Good | 20.14 | 54.90 | Regular |
| 14 | F | Poor | 23.25 | 54.00 | Infrequent |
| 15 | M | Fair | 23.30 | 67.50 | Regular |
| 16 | M | Good | 26.03 | 80.00 | Regular |
| 17 | F | Poor | 23.55 | 60.30 | Occasional |
| 18 | M | Fair | 28.89 | 78.75 | Regular |
| 19 | F | Very Good | 20.97 | 60.75 | Regular |
| 20 | M | Poor | 25.63 | 72.00 | Infrequent |
| 21 | F | Poor | 21.97 | 56.25 | Regular |
| 22 | M | Poor | 22.91 | 78.75 | Occasional |
| 23 | M | Fair | 25.02 | 83.70 | Regular |
| 24 | F | Excellent | 22.32 | 53.55 | Infrequent |
| 25 | M | Good | 27.16 | 81.00 | Regular |
| 26 | F | Fair | 23.66 | 62.55 | Regular |
| 27 | M | Poor | 24.70 | 94.50 | Infrequent |
| 28 | M | Fair | 21.97 | 67.50 | Infrequent |
| 29 | F | Very Good | 23.44 | 56.25 | Infrequent |
| 30 | F | Poor | 23.44 | 63.90 | Infrequent |
| 31 | F | Good | 31.37 | 85.50 | Missing |
| 32 | M | Fair | 33.40 | 96.75 | Regular |
| 33 | M | Poor | 27.69 | 90.00 | Regular |
| 34 | F | Good | 24.83 | 69.75 | Regular |
| 35 | F | Good | 31.75 | 78.75 | Infrequent |
| 36 | M | Poor | 22.63 | 67.50 | Infrequent |
| 37 | F | Fair | 24.47 | 53.10 | Regular |
| 38 | F | Good | 20.64 | 56.25 | Regular |
| 39 | M | Poor | 25.53 | 67.50 | Regular |
| 40 | F | Fair | 22.43 | 63.00 | Infrequent |
| 41 | F | Very Good | 26.03 | 58.50 | Regular |
| 42 | M | Poor | 22.06 | 69.75 | Infrequent |
| 43 | F | Poor | 25.49 | 65.25 | Regular |
| 44 | M | Poor | 22.28 | 78.75 | Occasional |
| 45 | F | Fair | 32.65 | 81.00 | Regular |
| 46 | F | Excellent | 24.55 | 79.81 | Regular |
| 47 | F | Good | 23.94 | 65.25 | Regular |
| 48 | M | Fair | 26.30 | 85.50 | Regular |
| 49 | M | Poor | 24.31 | 60.30 | Regular |
| 50 | M | Fair | 28.74 | 83.25 | Regular |
| 51 | M | Very Good | 29.79 | 78.75 | Regular |
| 52 | F | Poor | 17.99 | 49.05 | Infrequent |
| 53 | F | Good | 21.75 | 63.00 | Infrequent |
| 54 | F | Fair | 23.73 | 60.75 | Occasional |
| 55 | F | Poor | 17.09 | 49.50 | Regular |

**Directions for Assignment:**

1. Create a SPSS dataset using the data provided above. Include variable names, labels, and values for missing data where appropriate. Your dataset should include 6 variables and 55 cases. If data is presented in alphanumeric form you must use this form here. Therefore, your data should be entered as displayed above. Make sure that all the cells are properly completed when creating the variables. (18 points)
2. Develop numeric codes for all string variables but retain the alphanumeric characteristics as value labels (recode into different variable). – Don’t forget to complete the information I the “variable view”. (9 points)

***At this point, save a copy of the dataset (Assignment 1 – Questions 1&2 – Name) and the output file (Assignment 1 – Questions 1&2 – Name) and upload these to the Assignment Dropbox. – After this close the Output file.***

1. Create a “Codebook” illustrating the Variable name, Variable, and any coding instructions. Use the sample I have provided as an example for formatting and set up. Make sure this is neat and well organized. Include both original variables and recoded variables (Create in MS Word) (13 Points)

***Save the Codebook from MS Word and name the file (Assignment 1 – Question 3 – Name) and upload to the Assignment Dropbox.***

1. Calculate the most appropriate measure of central tendency, dispersion, and graph/chart for each variable (remember that ID is not a variable, so we do not do statistics for it). Use the recoded variables and not the original variables. Complete a write up for each variable discussing the descriptive statistics in MS Word. Tell the story, don’t just give me numbers!!!!! (5 points for the Output, 10 points for the write up)

***Save the SPSS Output and name the file (Assignment 1 – Question 4 – Name). This output should only include the descriptive statistics and graphs for each variable. Upload to the Assignment Dropbox. – After this close the Output file.***

***Save the written portion of this question from MS Word and name the file (Assignment 1 – Question 4 – Name) and upload to the Assignment Dropbox.***

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | |
| Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error |
| Body Mass Index | 55 | 17.09 | 37.62 | 24.7191 | 3.96456 | .860 | .322 |
| Weight (kg) | 55 | 43.65 | 96.75 | 70.5284 | 13.26490 | .190 | .322 |
| Valid N (listwise) | 55 |  |  |  |  |  |  |

The number of observations is 55 for both Body Mass Index (BMI) and Weight and there is no missing value. The mean of BMI and Weight is 24.7191 and 13.2649, respectively.

The skewness statistics value of BMI and Weight variable is 0.860 and 0.190, respectively, which indicates that both the variables are moderately skewed. In other words, the BMI and Weight variable follows normal distribution as shown below figure.

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

1. Using SPSS, calculate the mean, median, and standard deviation for “Weight” for all cases. Next, re-compute based on gender (use the split-file option). Complete a brief write up discussing any differences and what happened to the output when you split the file according to gender. Therefore, compare the original variable’s descriptive statistics to the variable after it is split by gender. The numbers should tell a story! Tell me what happened, DO NOT simply report the numbers. (5 points for the SPSS Output, 5 points for the write up)

***Save the SPSS Output and name the file (Assignment 1 – Question 5 – Name). The output file will include a description of the split file function and your descriptive statistics for the entire dataset, and also for the split dataset. Upload to the Assignment Dropbox. – After this close the Output file.***

***Save the written portion of this question from MS Word and name the file (Assignment 1 – Question 5 – Name) and upload to the Assignment Dropbox.***

1. Find the mean, median, and standard deviation of “Weight”. Then, using the filter option in SPSS, filter out all respondents that claim that their frequency of physical activity is “infrequent”. Re-run the descriptive statistics for “Weight”. Discuss the differences and what may have happened to weight, once you removed respondents who performed any type of physical actively infrequently. (5 points for the Output, 5 points for the write up)

***Save the SPSS Output and name the file (Assignment 1 – Question 6 – Name). The output file will include a description of the filter function and your descriptive statistics for the entire dataset, and also for the filtered dataset. Upload to the Assignment Dropbox. – After this close the Output file.***

***Save the written portion of this question from MS Word and name the file (Assignment 1 – Question 6 – Name) and upload to the Assignment Dropbox.***

Things to remember:

* This assignment is due on October 5, 2021 at 11:59 pm.
* Late assignments will not be accepted.
* This assignment is out of 75 points, and worth 15% of your final grade.
* Make sure you take the “filter option” and the “split-file option” off when you are not using it.
* It is not enough to just write exactly what the output reads. You need to fully explain each point you make and tell the story both in statistical and lay-person terms.
* Double check your data when you enter it. A typing error will distort all your outputs. (use copy and paste)
* Do not plagiarize. Plagiarism includes working with someone and handing in the same work. This is an independent assignment and not a group assignment.
* If you have any questions email me or arrange a time to meet with the instructor.
* DO NOT LEAVE THIS TO THE LAST MINUTE!