/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*

\* Project : BIOS 645 Course

\*

\* Program name : Homework 8

\*

\* Author : Esther Mun

\*

\* Date created : 2021-05-01

\*

\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

ods rtf file = '/home/u41545557/Bios511/Output/Homework8.rtf';

/\*

Imports the txt file into a work dataset

\*/

proc import datafile = '/home/u41545557/Bios511/Data/Quiz\_questions2.txt'

out = quiz

dbms = dlm

replace;

getnames=yes;

delimiter = '09'x;

run;

/\* First, conduct descriptive statistics on the variables to look at the data. \*/

\*Look at the overall dataset;

/\* proc print data = quiz; \*/

/\* run; \*/

/\* Question 13: Assuming that 0 means the student got the question wrong and

1 means the student got the question correct, 72% of students answered correctly and 14% of students answered incorrectly.

Over the semester, the minimum number of correct answers a student got is 24 and the maximum number of correct answers a student got is 49.

The median numer of correct answers was 37 and the mean was 37.6.

Since the mean is slightly greater than the median, the distribution will be slightly skewed.

The standard deviation is 7.106. \*/

PROC FREQ DATA=quiz;

TABLES q13\_y;

run;

PROC MEANS DATA=quiz N MIN MEDIAN MAX MEAN STD SKEW KURT MAXDEC=3;

VAR q13\_x;

run;

/\* Question 22:

Assuming that 0 means the student got the question wrong and

1 means the student got the question correct, 44% of students answered correctly and 56% of students answered incorrectly.

Over the semester, the minimum number of correct answers a student got is 24 and the maximum number of correct answers a student got is 49.

The median numer of correct answers was 38 and the mean was 37.88. Since the mean is slightly greater than the median, the distribution will be slightly skewed.

The standard deviation is 7.233.\*/

PROC FREQ DATA=quiz;

TABLES q22\_y;

run;

PROC MEANS DATA=quiz N MIN MEDIAN MAX MEAN STD SKEW KURT MAXDEC=3;

VAR q22\_x;

run;

/\* Question 41: Assuming that 0 means the student got the question wrong and

1 means the student got the question correct, 24% of students answered correctly and 76% of students answered incorrectly.

Over the semester, the minimum number of correct answers a student got is 24 and the maximum number of correct answers a student got is 49.

The median numer of correct answers was 38 and the mean was 38.080. Since the mean is slightly greater than the median, the distribution will be slightly skewed.

The standard deviation is 6.975. \*/

PROC FREQ DATA=quiz;

TABLES q41\_y;

run;

PROC MEANS DATA=quiz N MIN MEDIAN MAX MEAN STD SKEW KURT MAXDEC=3;

VAR q41\_x;

run;

/\* I regressed whether the student got the specific question right over the total number of correct answers the student got over the course. \*/

/\* This regresses whether the student got question 13 correct over the total number of correct answers they got

(besides question 13) over the course of the semester\*/

PROC LOGISTIC DATA=quiz DESCENDING PLOTS(ONLY)=EFFECT;

MODEL q13\_y=q13\_x;

\* the residual plots are hard to interpret, so we use "(ONLY)";

RUN;

/\* This regresses whether the student got question 22 correct over the total number of correct answers they got

(besides question 22) over the course of the semester\*/

PROC LOGISTIC DATA=quiz DESCENDING PLOTS(ONLY)=EFFECT;

MODEL q22\_y=q22\_x;

\* the residual plots are hard to interpret, so we use "(ONLY)";

RUN;

/\* This regresses whether the student got question 41 correct over the total number of correct answers they got

(besides question 41) over the course of the semester\*/

PROC LOGISTIC DATA=quiz DESCENDING PLOTS(ONLY)=EFFECT;

MODEL q41\_y=q41\_x;

\* the residual plots are hard to interpret, so we use "(ONLY)";

RUN;

ods rtf close;