/\* (a) import data "Project1Data.csv" \*/

**proc** **import** out=proj1data

datafile="C:/Users/Olga Korosteleva/Desktop/Project1Data.csv"

dbms=csv replace;

**run**;

**proc** **contents** varnum;

**run**;

/\* Answer the following questions: \*/

/\* (b) How many patients are there? \*/

**proc** **means** n;

var patientID;

**run**;

/\* (c) Compute age for each patient. \*/

**data** proj1data;

set proj1data;

today=**'12NOV2020'd**;

age=(today-DOB)/**365.25**;

**run**;

**proc** **print**;

var patientID age;

**run**;

/\* (d) Compute length of stay for each patient. \*/

**data** proj1data;

set proj1data;

length=date\_discharge-date\_admission;

**run**;

**proc** **print**;

var patientID length;

**run**;

/\* (e) Compute frequencies for gender. \*/

**proc** **freq**;

table gender/nopercent nocum;

**run**;

/\* (f) Compute frequencies for insurance types. \*/

**proc** **freq**;

table insurance/nopercent nocum;

**run**;

/\* (g) Compute basic descriptives for payment. \*/

**proc** **means** n mean median mode Q1 Q3 var std min max;

var payment;

**run**;

/\* (h) Compute frequencies of recovery program

(biogel - yes/no) by surgery site. \*/

**proc** **freq**;

table biogel\*site/norow nocol nopercent;

**run**;

/\* (i) Compute BMI. \*/

**data** proj1data;

set proj1data;

BMI=weight/height\*\***2**\***10000**;

**run**;

**proc** **print**;

var patientID BMI;

**run**;

/\* (j) Categorize each patient into

Underweight (BMI is less than 18.5), Normal (BMI

is 18.5 to 24.9), Overweight (BMI is 25 to 29.9),

and Obese (BMI is 30 or more) and compute frequencies.\*/

**data** proj1data;

set proj1data;

if (BMI<**18.5**) then BMI\_cat='underweight';

if (BMI>=**18.5** and BMI<**25**) then BMI\_cat='normal';

if (BMI>=**25** and BMI<**30**) then BMI\_cat='overweight';

if (BMI>=**30**) then BMI\_cat='obese';

**run**;

**proc** **print**;

var patientID BMI BMI\_cat;

**run**;