/\* Step 1: Create the health\_survey dataset \*/

data health\_survey;

do ID = 1 to 200;

Gender = ifc(ranuni(0) < 0.5, 'Male', 'Female');

Age = int(18 + ranuni(0)\*50);

BMI = round(18 + ranuni(0)\*17, 0.1);

Smoking\_Status = ifc(ranuni(0) < 0.3, 'Smoker', 'Non-Smoker');

Exercise\_Freq = int(ranuni(0)\*7);

Cholesterol = round(150 + ranuni(0)\*100, 1);

Diabetes = ifc(ranuni(0) < 0.2, 'Yes', 'No');

Blood\_Pressure = choose(ceil(ranuni(0)\*3), 'Low', 'Normal', 'High');

output;

end;

run;

/\* 1. Average BMI and Std Dev \*/

proc means data=health\_survey mean std;

var BMI;

run;

/\* 2. Min and Max Age by Gender \*/

proc summary data=health\_survey nway;

class Gender;

var Age;

output out=age\_summary min=MinAge max=MaxAge;

run;

proc print data=age\_summary; run;

/\* 3. Mean Cholesterol by Smoking\_Status \*/

proc means data=health\_survey mean;

class Smoking\_Status;

var Cholesterol;

run;

/\* 4. Frequency of Diabetes by Blood Pressure \*/

proc freq data=health\_survey;

tables Blood\_Pressure\*Diabetes / nocum nopercent;

run;

/\* 5. Frequency table for Smoking\_Status \*/

proc freq data=health\_survey;

tables Smoking\_Status;

run;

/\* 6. Contingency table between Gender and Diabetes \*/

proc freq data=health\_survey;

tables Gender\*Diabetes / chisq;

run;

/\* 7. Smoking Status vs Blood Pressure (with %s) \*/

proc freq data=health\_survey;

tables Smoking\_Status\*Blood\_Pressure / chisq norow nocol nopercent;

run;

/\* 8. T-Test: BMI by Gender \*/

proc ttest data=health\_survey;

class Gender;

var BMI;

run;

/\* 9. T-Test: Cholesterol by Smoking\_Status \*/

proc ttest data=health\_survey;

class Smoking\_Status;

var Cholesterol;

run;

/\* 10. Non-parametric test: Exercise\_Freq by Diabetes \*/

proc npar1way data=health\_survey wilcoxon;

class Diabetes;

var Exercise\_Freq;

run;

/\* 11. Chi-square: Diabetes vs Blood Pressure \*/

proc freq data=health\_survey;

tables Diabetes\*Blood\_Pressure / chisq;

run;

/\* 12. Correlation: Age, BMI, Cholesterol \*/

proc corr data=health\_survey;

var Age BMI Cholesterol;

run;

/\* 13. T-Test: Age by Diabetes \*/

proc ttest data=health\_survey;

class Diabetes;

var Age;

run;

/\* 14. Bar Chart: Gender \*/

proc sgplot data=health\_survey;

vbar Gender;

run;

/\* 15. Histogram of BMI \*/

proc sgplot data=health\_survey;

histogram BMI;

run;

/\* 16. Scatter plot: BMI vs Cholesterol + regression line \*/

proc sgplot data=health\_survey;

scatter x=BMI y=Cholesterol;

reg x=BMI y=Cholesterol;

run;

/\* 17. Box plot: Cholesterol by Blood Pressure \*/

proc sgplot data=health\_survey;

vbox Cholesterol / category=Blood\_Pressure;

run;

/\* 18. Grouped bar chart: Smoking Status by Gender \*/

proc sgplot data=health\_survey;

vbar Smoking\_Status / group=Gender groupdisplay=cluster;

run;

/\* 19. Summary stats: Exercise\_Freq by Smoking\_Status & Gender \*/

proc summary data=health\_survey nway;

class Smoking\_Status Gender;

var Exercise\_Freq;

output out=exercise\_summary mean= mean median= median std= stddev;

run;

proc print data=exercise\_summary; run;

/\* 20. Summary report by Diabetes status \*/

proc report data=health\_survey nowd;

column Diabetes Age BMI Cholesterol;

define Diabetes / group;

define Age / analysis mean;

define BMI / analysis mean;

define Cholesterol / analysis mean;

run;