**Module 8 Portfolio Project Code**

Brooke Dietrich

Github Link and Project Code

MIS581: Capstone: Business Intelligence and Data Analytics

Dr. Steve Chung

May 14, 2023

**Project Code**

**Importing the Netflix Dataset onto SAS Studio**

/\* (IMPORT) \*/

/\* Source File: Netflix Dataset.csv \*/

/\* Source Path: /home/u60844558/sasuser.v94 \*/

/\* Code generated on: 5/13/23, 1:59 PM \*/

%if %sysfunc(exist(WORK.'Netflix Dataset1'n)) %then %do;

proc sql;

drop table WORK.'Netflix Dataset1'n;

run;

%end;

FILENAME REFFILE '/home/u60844558/sasuser.v94/Netflix Dataset.csv';

PROC IMPORT DATAFILE=REFFILE

DBMS=CSV

OUT=WORK.'Netflix Dataset1'n;

GETNAMES=YES;

RUN;

PROC CONTENTS DATA=WORK.'Netflix Dataset1'n; RUN;

%web\_open\_table(WORK.'Netflix Dataset1'n);

**Creating a Bar Graph in SAS Studio**

proc sgplot data=WORK.'NETFLIX DATASET1'n;

vbar type /;

yaxis grid;

run;

Chart, bar chart

Description automatically generated

proc sgplot data=WORK.'NETFLIX DATASET1'n;

vbar rating /;

yaxis grid;

run;

*Chart, histogram

Description automatically generated*

**Creating a Pie Chart in SAS Studio**

proc template;

define statgraph SASStudio.Pie;

begingraph;

layout region;

piechart category=type /;

endlayout;

endgraph;

end;

run;

Chart, pie chart

Description automatically generated

**Correlation between Release Year and Date Added in Python**

**A screenshot of a computer

Description automatically generated**

**R-Square Value for Release Year versus Date Added**

proc glmselect data=WORK.'NETFLIX DATASET1'n

outdesign(addinputvars)=Work.reg\_design;

class release\_year / param=glm;

model date\_added=release\_year / showpvalues selection=none;

run;

proc reg data=Work.reg\_design alpha=0.05 plots(only)=(diagnostics residuals

observedbypredicted);

where release\_year is not missing;

ods select DiagnosticsPanel ResidualPlot ObservedByPredicted;

model date\_added=&\_GLSMOD /;

run;

*Table

Description automatically generated*

**R-Square Value for Release Year vs Rating**

proc glmselect data=WORK.'NETFLIX DATASET1'n

outdesign(addinputvars)=Work.reg\_design;

class rating / param=glm;

model release\_year=rating / showpvalues selection=none;

run;

proc reg data=Work.reg\_design alpha=0.05 plots(only)=(diagnostics residuals

observedbypredicted);

where rating is not missing;

ods select DiagnosticsPanel ResidualPlot ObservedByPredicted;

model release\_year=&\_GLSMOD /;

run;

quit;

proc delete data=Work.reg\_design;

run;

*Table

Description automatically generated*

**T-Test for Date Added**

proc univariate data=WORK.'NETFLIX DATASET1'n normal mu0=0;

ods select TestsForNormality;

var date\_added;

run;

/\* t test \*/

proc ttest data=WORK.'NETFLIX DATASET1'n sides=2 h0=0 plots(showh0);

var date\_added;

run;

*Graphical user interface, application

Description automatically generated*

Github link: <https://github.com/brookehd12/NetflixProject>

Youtube video link: <https://youtu.be/nQOCCw62Btk>