**Module 8 Portfolio Project Code**

**Joseph Yevara**

**Colorado State University Global**

# **MIS581-1: Capstone: Business Intelligence and Data Analytics**

# **Dr. Osama Morad**

**02/11/2024**

**Socioeconomic Factors**

**Descriptive Statistics**

83 proc freq data=WIP.SES;

84 tables 'Marital status'n 'Application mode'n 'Application order'n

85 'Previous qualification'n Nacionality 'Mother''s qualification'n

86 'Father''s qualification'n 'Mother''s occupation'n 'Father''s occupation'n

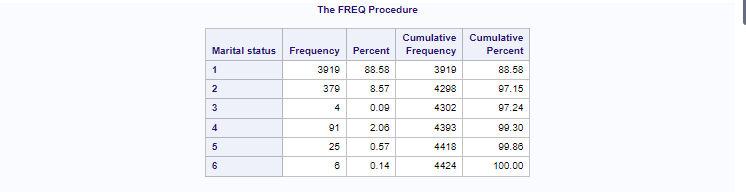
87 Displaced 'Educational special needs'n Debtor 'Tuition fees up to date'n

88 Gender 'Scholarship holder'n International / plots=(freqplot cumfreqplot);

89 run;

**Figure 1**

*Partial output of one-way frequencies*



Note: Each variable had an output like this that was used to create the table included within the final project

**Regression Analysis**

87 proc logistic data=WIP.SES;

88 class 'Marital status'n 'Application mode'n 'Application order'n Nacionality

89 'Mother''s qualification'n 'Father''s qualification'n 'Mother''s occupation'n

90 'Father''s occupation'n Displaced 'Educational special needs'n Debtor

91 'Tuition fees up to date'n Gender 'Scholarship holder'n International /

92 param=glm;

93 model Target(event='Graduate')='Marital status'n 'Application mode'n

94 'Application order'n Nacionality 'Mother''s qualification'n

95 'Father''s qualification'n 'Mother''s occupation'n 'Father''s occupation'n

96 Displaced 'Educational special needs'n Debtor 'Tuition fees up to date'n

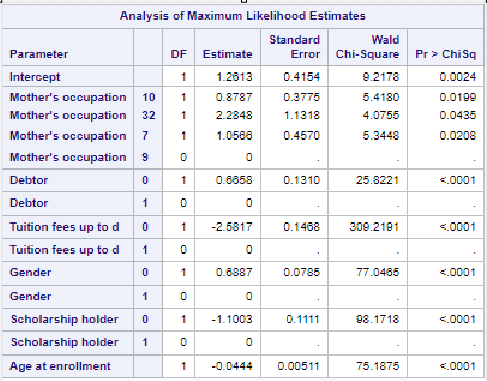
97 Gender 'Scholarship holder'n International 'Age at enrollment'n / link=logit

98 selection=stepwise slentry=0.05 slstay=0.05 hierarchy=single technique=fisher;

99 run;

**Figure 2**

*Partial output of logistic regression*

****

**Academic Factors**

**Descriptive Statistics**

87 proc means data=WIP.ACADEMIC chartype mean std min max n vardef=df;

88 var 'Curricular units 1st sem (credit'n 'Curricular units 1st sem (enroll'n

89 'Curricular units 1st sem (evalua'n 'Curricular units 1st sem (approv'n

90 'Curricular units 1st sem (grade)'n 'Curricular units 1st sem (withou'n

91 'Curricular units 2nd sem (credit'n 'Curricular units 2nd sem (enroll'n

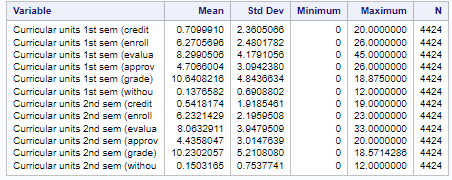
92 'Curricular units 2nd sem (evalua'n 'Curricular units 2nd sem (approv'n

93 'Curricular units 2nd sem (grade)'n 'Curricular units 2nd sem (withou'n;

94 run;

**Figure 3**

*Summary statistics output*



**Correlation Analysis**

87 proc corr data=WIP.ACADEMIC pearson nosimple noprob plots=none;

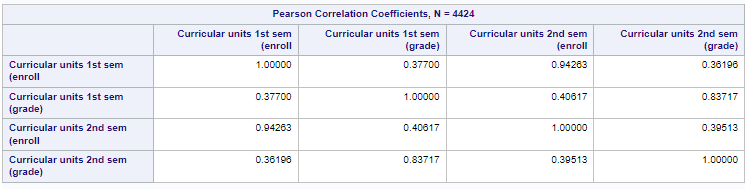
88 var 'Curricular units 1st sem (enroll'n 'Curricular units 1st sem (grade)'n

89 'Curricular units 2nd sem (enroll'n 'Curricular units 2nd sem (grade)'n;

90 run;

**Figure 4**

*Correlation analysis output*



**Regression Analysis**

87 proc logistic data=WIP.ACADEMIC;

88 class Course 'Daytime/evening attendance'n / param=glm;

89 model Target(event='Graduate')=Course 'Daytime/evening attendance'n

90 'Curricular units 2nd sem (grade)'n 'Curricular units 2nd sem (enroll'n /

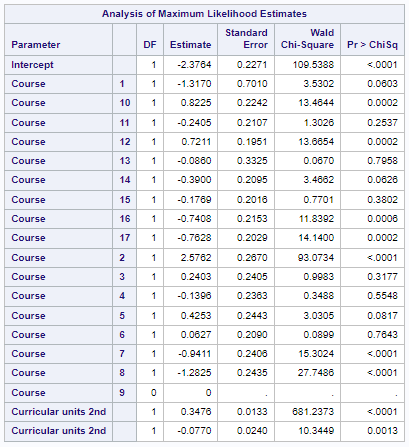
91 link=logit selection=stepwise slentry=0.05 slstay=0.05 hierarchy=single

92 technique=fisher;

93 run;

**Figure 5**

*Partial output of logistic regression*



**Economic Factors**

**Descriptive Statistics**

87 proc means data=WIP.ECON chartype mean std min max n vardef=df;

88 var 'Unemployment rate'n 'Inflation rate'n GDP;

89 run;

90

91 /\* Graph template to construct combination histogram/boxplot \*/

92 proc template;

93 define statgraph histobox;

94 dynamic AVAR ByVarInfo;

95 begingraph;

96 entrytitle "Distribution of " AVAR ByVarInfo;

97 layout lattice / rows=2 columndatarange=union rowgutter=0 rowweights=(0.75

98 0.25);

99 layout overlay / yaxisopts=(offsetmax=0.1) xaxisopts=(display=none);

100 histogram AVAR /;

101 endlayout;

102 layout overlay /;

103 BoxPlot Y=AVAR / orient=horizontal;

104 endlayout;

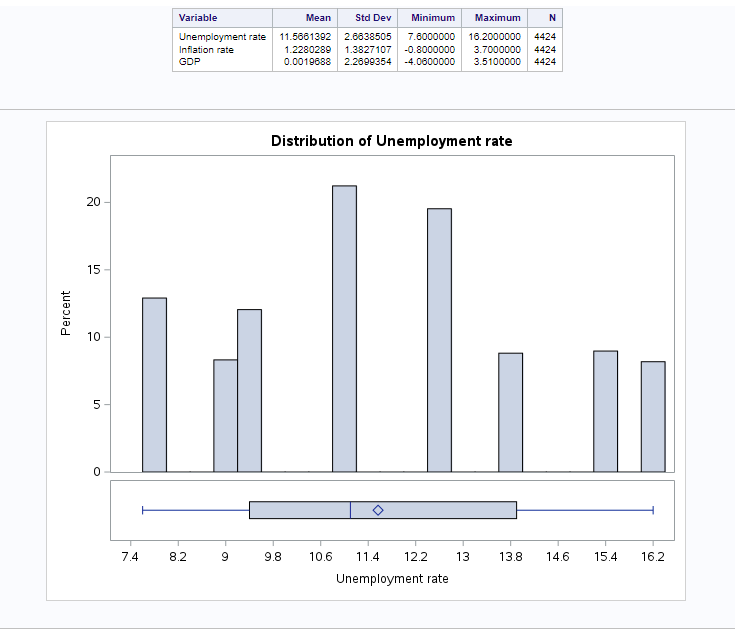
105 endlayout;

106 endgraph;

107 end;

**Figure 6**

*Partial output of summary statistics*



Note: Each variable boxplot was included in paper but not histogram.

**Correlation Analysis**

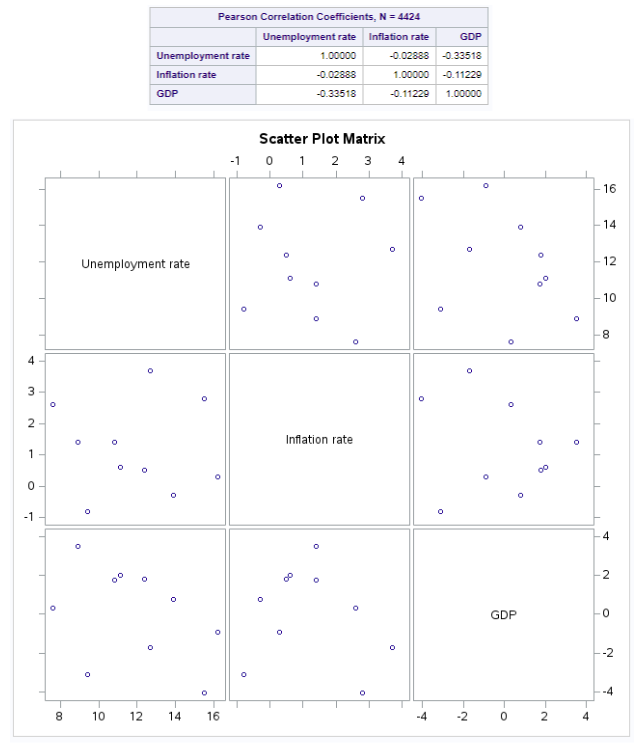
87 proc corr data=WIP.ECON pearson nosimple noprob plots(maxpoints=none)=matrix;

88 var 'Unemployment rate'n 'Inflation rate'n GDP;

89 run;

**Figure 7**

*Correlation analysis output*



**Regression Analysis**

87 proc logistic data=WIP.ECON;

88 model Target(event='Graduate')='Unemployment rate'n 'Inflation rate'n GDP /

89 link=logit selection=stepwise slentry=0.05 slstay=0.05 hierarchy=single

90 technique=fisher;

91 run;

**Figure 8**

*Partial output of regression analysis*

