

# Syndicate 5 Statistical Learning Problem Set #1

May 18, 2021

Table 1: Correlation Matrix (numeric data only)

	Income	Limit	Rating	Cards	Age	Balance
Income	1	0.792	0.791	-0.018	0.175	0.464
Limit	0.792	1	0.997	0.010	0.101	0.862
Rating	0.791	0.997	1	0.053	0.103	0.864
Cards	-0.018	0.010	0.053	1	0.043	0.086
Age	0.175	0.101	0.103	0.043	1	0.002
Balance	0.464	0.862	0.864	0.086	0.002	1

Table 2: Model Selection

	<i>Dependent variable:</i>		
	Balance		
	(1)	(2)	(3)
Income	−6.550*** (0.445)	−6.610*** (0.447)	−6.350*** (0.489)
I(Income^2)	−0.020*** (0.003)	−0.020*** (0.003)	−0.021*** (0.003)
Limit	0.014 (0.053)	0.027 (0.053)	
I(Limit^2)	0.00002*** (0.00000)	0.00002*** (0.00000)	
Rating	2.540*** (0.867)	2.350*** (0.862)	2.490*** (0.137)
I(Rating^2)	−0.002** (0.001)	−0.002** (0.001)	0.002*** (0.0002)
Cards	25.800** (10.900)	17.900*** (3.550)	2.520 (3.270)
I(Cards^2)	−1.110 (1.440)		
Age	0.254 (1.600)		
I(Age^2)	−0.009 (0.014)		
Edu_BinsBachelors	7.000 (9.760)		
Edu_BinsPost-Grad	−15.000 (10.000)		
GenderFemale	−8.260 (8.080)		
StudentYes	442.000*** (13.700)	441.000*** (13.600)	429.000*** (14.900)
MarriedYes	−6.130 (8.430)		
EthnicityAsian	16.900 (11.500)	19.000* (11.500)	21.300* (12.600)