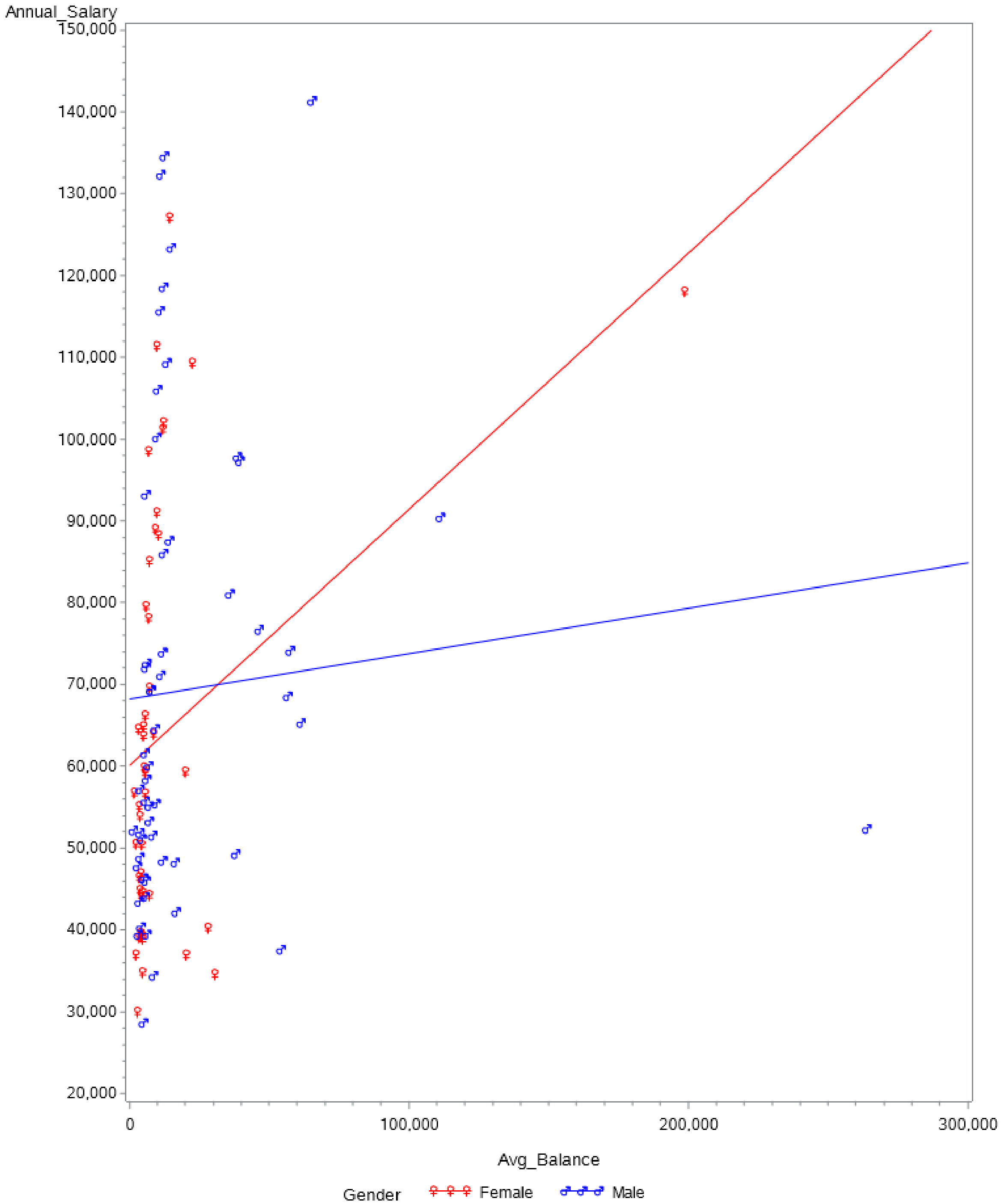
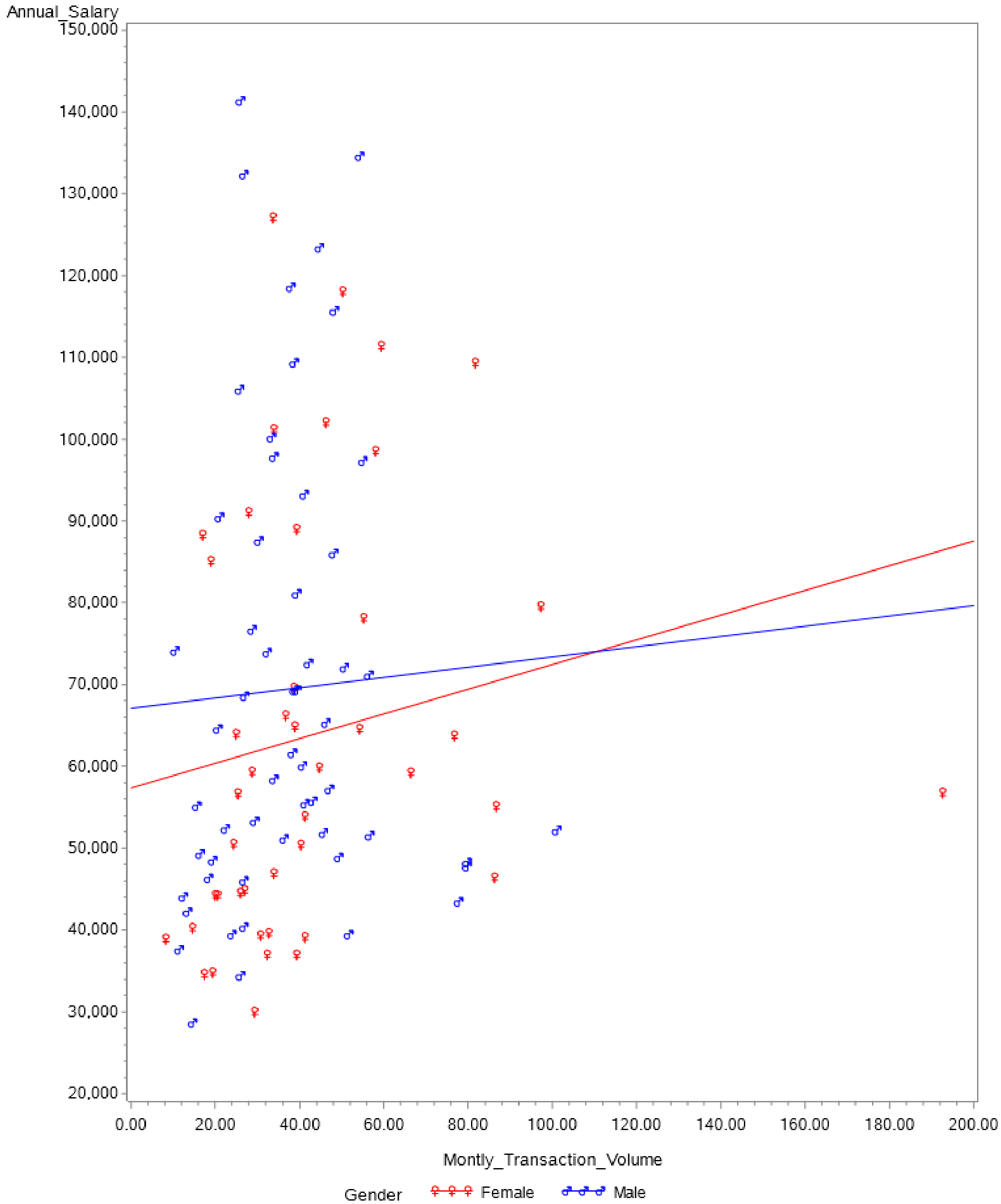
Module 2 Data @ANZ

1. Annual Salary Scatterplots

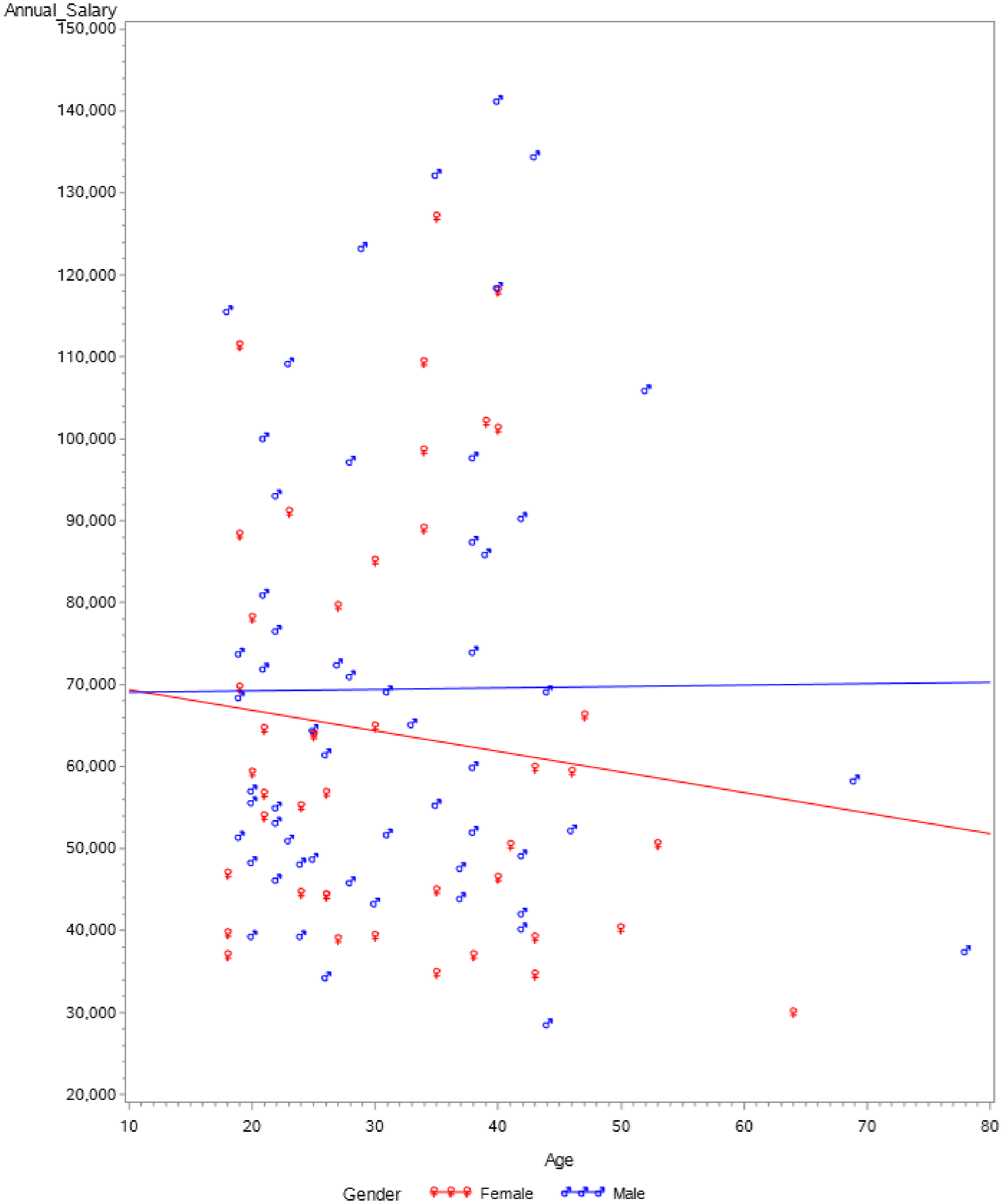
Using SAS, I created a scatterplot to compare Annual Salary with Average Balance, Monthly Transaction Volume, and Age. I separated the data set by gender and created simple regression lines. Interestingly, Annual Salary goes up more for females as Average Balance goes up. On the other hand, Annual Salary lowers more for females as Age goes up. Below each graph is the simple regression equation.



Regression Equation:
Annual_Salary(Gender:Female) = 60115.57 + 0.313271*Avg_Balance
Annual_Salary(Gender:Male) = 68291.29 + 0.055155*Avg_Balance



Regression Equation:
Annual_Salary(Gender:Female) = 57387.97 + 150.8618*Montly_Transaction_Volume
Annual_Salary(Gender:Male) = 67150.44 + 62.65279*Montly_Transaction_Volume



Regression Equation:
Annual_Salary(Gender:Female) = 71916.04 - 251.3882*Age
Annual_Salary(Gender:Male) = 68933.5 + 17.58486*Age

2. Parameter Estimates

After doing a Standard Least Squares Model, we see only Avg_Balance has significant evidence to reject the null hypothesis at 95% confidence.

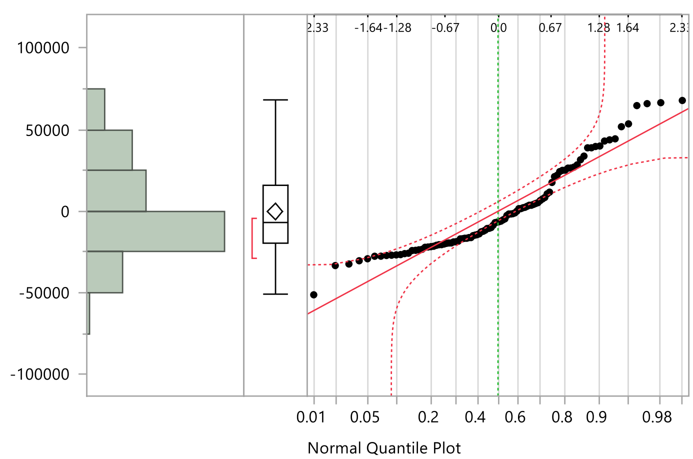
Response Annual_Salary

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	63170.971	9707.719	6.51	<.0001*
Gender[Female]	-2464.802	2726.395	-0.90	0.3683
Age	-149.755	241.0783	-0.62	0.5360
Avg_Balance	0.1577108	0.078837	2.00	0.0483*
Montly_Transaction_Volume	138.85583	110.7085	1.25	0.2128

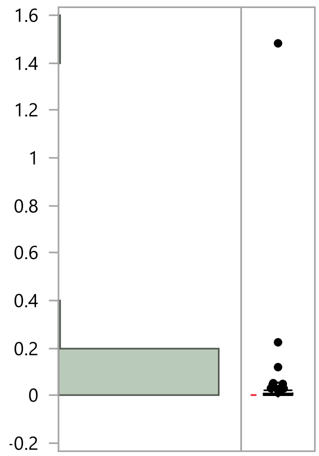
3. Model Assumptions

Looking at the residuals and Cook D's, our model assumptions are wrong. The figure below shows our residuals are not all within the bands of homoscedasticity telling us our model is not homogenous. Cook D's Influence tells us there is one outlier (Patrick) that significantly changes our model.

Residual Annual_Salary



Cook's D Influence Annual_Salary



4. Prediction Profiler

If we disregard the model assumptions, we can predict the annual salary our customers using the prediction profiler below. We can see being male, young, and having high average balance and transaction volume gives us a higher predicted annual salary for our customers.

Prediction Profiler