**Quantitative Analyst Coding Interview**

**Instructions for Candidates:**

This interview is designed to evaluate your approach to solving coding related problems. You will be shown a series of problems and asked how you would go about creating code to solve them. Please write out a sketch of the code that you would write. The questions will be based on problems that would be solved in a real-world setting. However, the coding involved does not need to rise to the level of working code. Pseudo code or outlines of coding logic are perfectly acceptable. If you would prefer to write your solutions with a specific language in mind, feel free to do so. However, you may also write it in a language agnostic fashion if you would prefer.

For example, if the question were: How would you find the average balance per account for accounts that have a credit limit under $1000

An example of an acceptable answer would be:

Filter accounts with credit\_limit < 1000

Group by account\_id

avg\_bal = mean(balance)

At Capital One we work with very large data, for example millions of observations and hundreds of columns. Wherever possible, try to design solutions that would be scalable to large datasets and would generalize to broader applications.

For this virtual interview, you will be providing your solutions in this document while sharing your screen so that the interviewer can follow along.

**Data Description**

The data used in these problems is based on Auto loans. Capital One has a large portfolio of Auto loans and frequently wants to monitor how these loans are performing and predict the outcomes for different loans. In this case there are 3 datasets involved:

1. Driver Table

This data contains internal performance data on each loan, such as current balance and payment history as well as some information on the loan, such as original loan amount and original vehicle value. The data is monthly for each account, beginning in month that the account was opened and continuing until the latest month of available data (Dec 2019 in this case)

1. Vehicle Value

This data contains estimates of every vehicle’s value for any given month. It is available at the make, model, year level and has observations for every month since the vehicle was introduced until the current month.

1. VIN

This data links each VIN to the vehicle’s manufacturing information.

**Solutions**