# DATA 624:  Project 1

*Vinicio Haro*
*Sang Yoon (Andy) Hwang*
*Julian McEachern*
*Jeremy O'Brien*
*Bethany Poulin*

*October 22, 2019*

# Contents

# Overview

> Insert Project Overview and explain work process.

## Dependencies

> Please add all libraries used here.

The following R libraries were used to complete Project 1:

```r
# Insert All Used Dependencies Here
```

## Data

Data was stored within our group repository and imported below using the `readxl` package. Each individual question was solved within an R script and the data was sourced into our main report for discussion purposes. The R scripts are available within our appendix for replication purposes.

For grading purposes, we exported and saved all forecasts as a csv in our data folder.

```r
# Data Aquisition
atm_data <- read_excel("data/ATM624Data.xlsx")
power_data <- read_excel("data/ResidentialCustomerForecastLoad-624.xlsx")
pipe1_data <- read_excel("data/Waterflow_Pipe1.xlsx")
pipe2_data <- read_excel("data/Waterflow_Pipe2.xlsx")

# Source Code
source("scripts/Part-A-JM.R") ## Example
```

# 1 Part A: ATMs

**Instructions:** In part A, I want you to forecast how much cash is taken out of 4 different ATM machines for May 2010. The data is given in a single file. The variable Cash is provided in hundreds of dollars, other than that it is straight forward. I am being somewhat ambiguous on purpose. I am giving you data, please provide your written report on your findings, visuals, discussion and your R code all within a Word readable document, except the forecast which you will put in an Excel readable file. I must be able to cut and paste your R code and run it in R studio. Your report must be professional - most of all - readable, EASY to follow. Let me know what you are thinking, assumptions you are making! Your forecast is a simple CSV or Excel file that MATCHES the format of the data I provide.

# 2 Part B: Power

**Instructions:** Forecasting Power: Part B consists of a simple dataset of residential power usage for January 1998 until December 2013. Your assignment is to model these data and a monthly forecast for 2014. The data is given in a single file. The variable 'KWH' is power consumption in Kilowatt hours, the rest is straight forward. Add these to your existing files above - clearly labeled.

# 3   Part C: Waterflow

**Instructions:** Part C consists of two data sets. These are simple 2 columns sets, however they have different time stamps. Your optional assignment is to time-base sequence the data and aggregate based on hour (example of what this looks like, follows). Note for multiple recordings within an hour, take the mean. Then to test appropriate assumptions and forecast a week forward with confidence bands (80 and 95%). Add these to your existing files above - clearly labeled.

# Appendix

## Part A

**ATM1 Forecast**

**ATM2 Forecast**

**ATM3 Forecast**

**ATM4 Forecast**

**R Script**

```
#Insert Script Here
```

## Part B

**Power Forecast**

**R Script**

```
#Insert Script Here
```

## Part C

**Pipes1 Forecast**

**Pipes2 Forecast**

**R Script**

```
#Insert Script Here
```