

Project 3 (first draft, in progress)

Author Name

2020-10-09

Data Loading

```
library(readxl)
library(httr)
```

Data Source (Excel file)

```
retailURL <- "http://archive.ics.uci.edu/ml/machine-learning-databases/retaildata/retail.xls"
GET(retailURL, write_disk(tempFileName <- tempfile(fileext = ".xls")))

## Response [http://archive.ics.uci.edu/ml/machine-learning-databases/retaildata/retail.xls]
##   Date: 2020-10-09 18:15
##   Status: 200
##   Content-Type: application/x-httpd-php
##   Size: 45.6 MB
## <ON DISK>  C:\Users\HannK\AppData\Local\Temp\RtmpoJTG32\tempfile00000000000000000000.xls

retail_sheet_2009 <- read_excel(tempFileName, sheet = "Year 2009")
retail_sheet_2010 <- read_excel(tempFileName, sheet = "Year 2010")
retaildf <- rbind(retail_sheet_2009, retail_sheet_2010)
```

Description of Data

```
head(retaildf)
```

```
## # A tibble: 6 x 8
```

```
##   Invoice StockCode Description Quantity InvoiceDate
```

```
##   <chr>    <chr>      <chr>          <dbl> <dtm>
```

```
## 1 489434  85048      "15CM CHRI~      12 2009-12-01 07:4
```

```
## 2 489434  79323P     "PINK CHER~      12 2009-12-01 07:4
```

```
## 3 489434  79323W     "WHITE CHE~      12 2009-12-01 07:4
```

```
## 4 489434  22041     "RECORD FR~      48 2009-12-01 07:4
```

```
## 5 489434  21232     "STRAWBERR~      24 2009-12-01 07:4
```

```
## 6 489434  22064     "PINK DOUG~      24 2009-12-01 07:4
```

```
## # ... with 1 more variable: Country <chr>
```

E-R Diagram

