

...

## Group 7

2022-08-24

Setup cran repository url

```
r = getOption("repos")
r["CRAN"] = "http://cran.us.r-project.org"
options(repos = r)
```

Install the required packages

```
#install.packages('readxl')
#install.packages('tidyverse')
#install.packages("lubridate")
#install.packages("dplyr")
#install.packages("imputeTS")
#install.packages("corrplot")
```

Set your current R working directory to the specific path.

```
path = dirname(rstudioapi::getSourceEditorContext())$path
setwd(path)
```

Read data from different files

```
data_months <- read.csv("data/dataset/months.csv")
cash_rate_data <- read.csv("data/dataset/Cash rate -RBA.csv")
loan_commitments <- read.csv("data/dataset/New loan commitments total housing.csv")
labour_force_data <- read_excel("data/dataset/Labour Force Australia.xlsx", sheet = "Data1")
```

FALSE New names:

```
FALSE * '' -> '...1'
FALSE * 'Employed total ; Persons ;' -> 'Employed total ; Persons ;...2'
FALSE * 'Employed total ; Persons ;' -> 'Employed total ; Persons ;...3'
FALSE * 'Employed total ; Persons ;' -> 'Employed total ; Persons ;...4'
FALSE * 'Employed total ; > Males ;' -> 'Employed total ; > Males ;...5'
FALSE * 'Employed total ; > Males ;' -> 'Employed total ; > Males ;...6'
FALSE * 'Employed total ; > Males ;' -> 'Employed total ; > Males ;...7'
FALSE * 'Employed total ; > Females ;' -> 'Employed total ; > Females ;...8'
FALSE * 'Employed total ; > Females ;' -> 'Employed total ; > Females ;...9'
FALSE * 'Employed total ; > Females ;' -> 'Employed total ; > Females ;...10'
FALSE * '> Employed full-time ; Persons ;' -> '> Employed full-time ; Persons ;...11'
FALSE * '> Employed full-time ; Persons ;' -> '> Employed full-time ; Persons ;...12'
```





```

FALSE * '> Unemployment rate looked for full-time work ; > Females ;' -> '>
FALSE   Unemployment rate looked for full-time work ; > Females ;...81'
FALSE * '> Unemployment rate looked for full-time work ; > Females ;' -> '>
FALSE   Unemployment rate looked for full-time work ; > Females ;...82'
FALSE * '> Unemployment rate looked for only part-time work ; Persons ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; Persons ;...83'
FALSE * '> Unemployment rate looked for only part-time work ; Persons ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; Persons ;...84'
FALSE * '> Unemployment rate looked for only part-time work ; Persons ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; Persons ;...85'
FALSE * '> Unemployment rate looked for only part-time work ; > Males ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Males ;...86'
FALSE * '> Unemployment rate looked for only part-time work ; > Males ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Males ;...87'
FALSE * '> Unemployment rate looked for only part-time work ; > Males ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Males ;...88'
FALSE * '> Unemployment rate looked for only part-time work ; > Females ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Females ;...89'
FALSE * '> Unemployment rate looked for only part-time work ; > Females ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Females ;...90'
FALSE * '> Unemployment rate looked for only part-time work ; > Females ;' -> '>
FALSE   Unemployment rate looked for only part-time work ; > Females ;...91'
FALSE * 'Labour force total ; Persons ;' -> 'Labour force total ; Persons ;...92'
FALSE * 'Labour force total ; Persons ;' -> 'Labour force total ; Persons ;...93'
FALSE * 'Labour force total ; Persons ;' -> 'Labour force total ; Persons ;...94'
FALSE * 'Labour force total ; > Males ;' -> 'Labour force total ; > Males ;...95'
FALSE * 'Labour force total ; > Males ;' -> 'Labour force total ; > Males ;...96'
FALSE * 'Labour force total ; > Males ;' -> 'Labour force total ; > Males ;...97'
FALSE * 'Labour force total ; > Females ;' -> 'Labour force total ; > Females ;...98'
FALSE * 'Labour force total ; > Females ;' -> 'Labour force total ; > Females ;...99'
FALSE * 'Labour force total ; > Females ;' -> 'Labour force total ; > Females
FALSE   ;...100'
FALSE * 'Participation rate ; Persons ;' -> 'Participation rate ; Persons ;...101'
FALSE * 'Participation rate ; Persons ;' -> 'Participation rate ; Persons ;...102'
FALSE * 'Participation rate ; Persons ;' -> 'Participation rate ; Persons ;...103'
FALSE * 'Participation rate ; > Males ;' -> 'Participation rate ; > Males ;...104'
FALSE * 'Participation rate ; > Males ;' -> 'Participation rate ; > Males ;...105'
FALSE * 'Participation rate ; > Males ;' -> 'Participation rate ; > Males ;...106'
FALSE * 'Participation rate ; > Females ;' -> 'Participation rate ; > Females
FALSE   ;...107'
FALSE * 'Participation rate ; > Females ;' -> 'Participation rate ; > Females
FALSE   ;...108'
FALSE * 'Participation rate ; > Females ;' -> 'Participation rate ; > Females
FALSE   ;...109'

```

```

dwellings_mean_price <- read_excel("data/dataset/Mean price of residential dwellings_Australia.xlsx", sl

```

```

FALSE New names:
FALSE * '' -> '...1'

```

```

residential_dwellings_number <- read_excel("data/dataset/Mean price of residential dwellings_Australia..

```

```

FALSE New names:
FALSE * '' -> '...1'

```

```
earnings <- read_excel("data/dataset/Average Weekly Earnings.xlsx", sheet = "Data1")
```

FALSE New names:

```
FALSE * '' -> '...1'
```

```
property_price_index <- read_excel("data/dataset/Residential Property Price Index.xlsx", sheet = "Data1")
```

FALSE New names:

```
FALSE * '' -> '...1'
```

You can view the object type

```
class(loan_commitments)
class(labour_force_data)
```

Retrieve relevant columns

```
loan_commitments <- loan_commitments[,c(1,2)]
labour_force_data <- labour_force_data[,c(1,67)]
dwellings_mean_price <- dwellings_mean_price[,c(1,37)]
residential_dwellings_number <- residential_dwellings_number[,c(1,46)]
earnings <- earnings[,c(1,9)]
property_price_index <- property_price_index[,c(1,10)]
```

Rename columns

Skip details rows

```
#labour_force_data <- labour_force_data[-c(1:9),]
#dwellings_mean_price <- dwellings_mean_price[-c(1:9),]
```

Change date format so it can be manipulated in R

```
str(data_months)
data_months$Date <- dmy(data_months$Date)
cash_rate_data$Date <- dmy(cash_rate_data$Date)
loan_commitments$Date <- dmy(loan_commitments$Date)
```

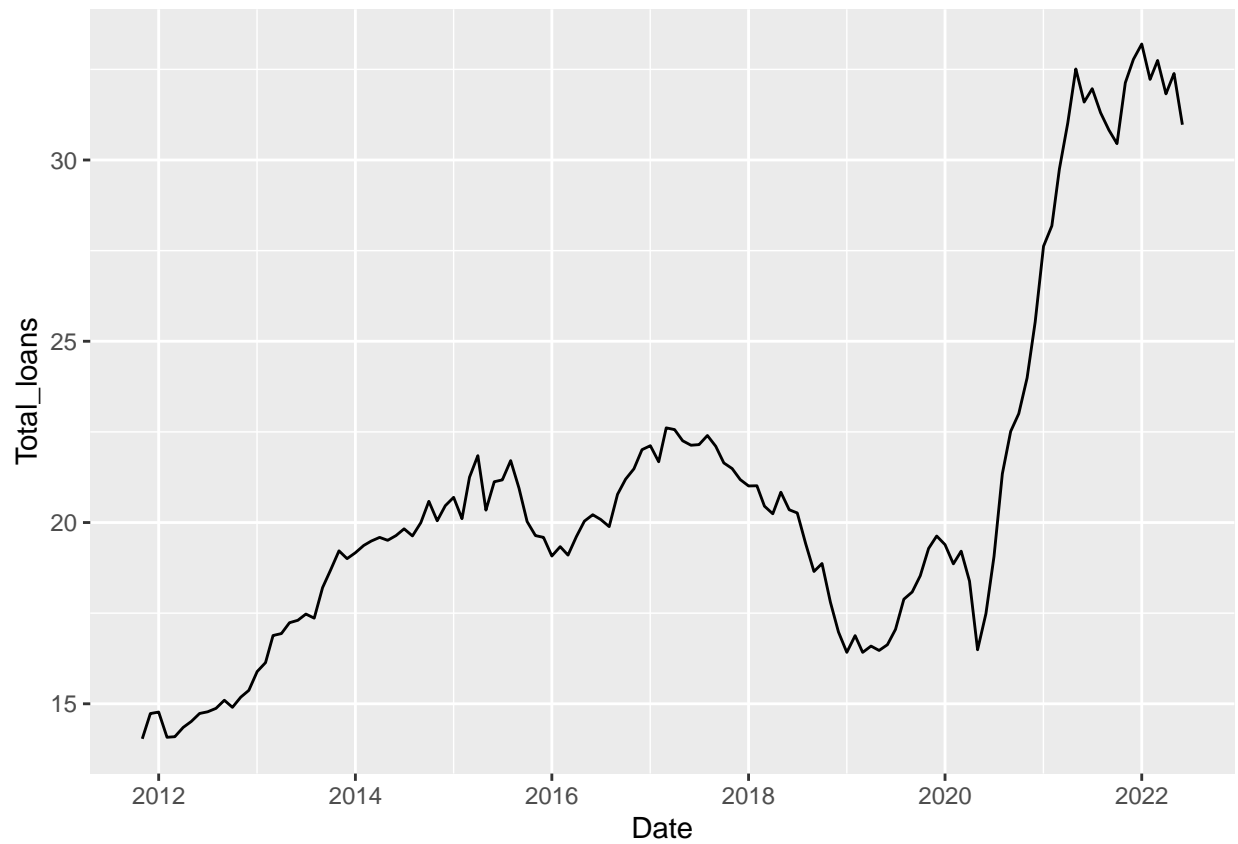
Create dataset by merging dataframes with dplyr

```
merge_dataset <- left_join(data_months, cash_rate_data, by='Date') %>%
  left_join(., labour_force_data, by='Date') %>%
  left_join(., dwellings_mean_price, by='Date') %>%
  left_join(., residential_dwellings_number, by='Date') %>%
  left_join(., earnings, by='Date') %>%
  left_join(., property_price_index, by='Date') %>%
  left_join(., loan_commitments, by='Date')
```

## UNDERSTANDING THE DATA

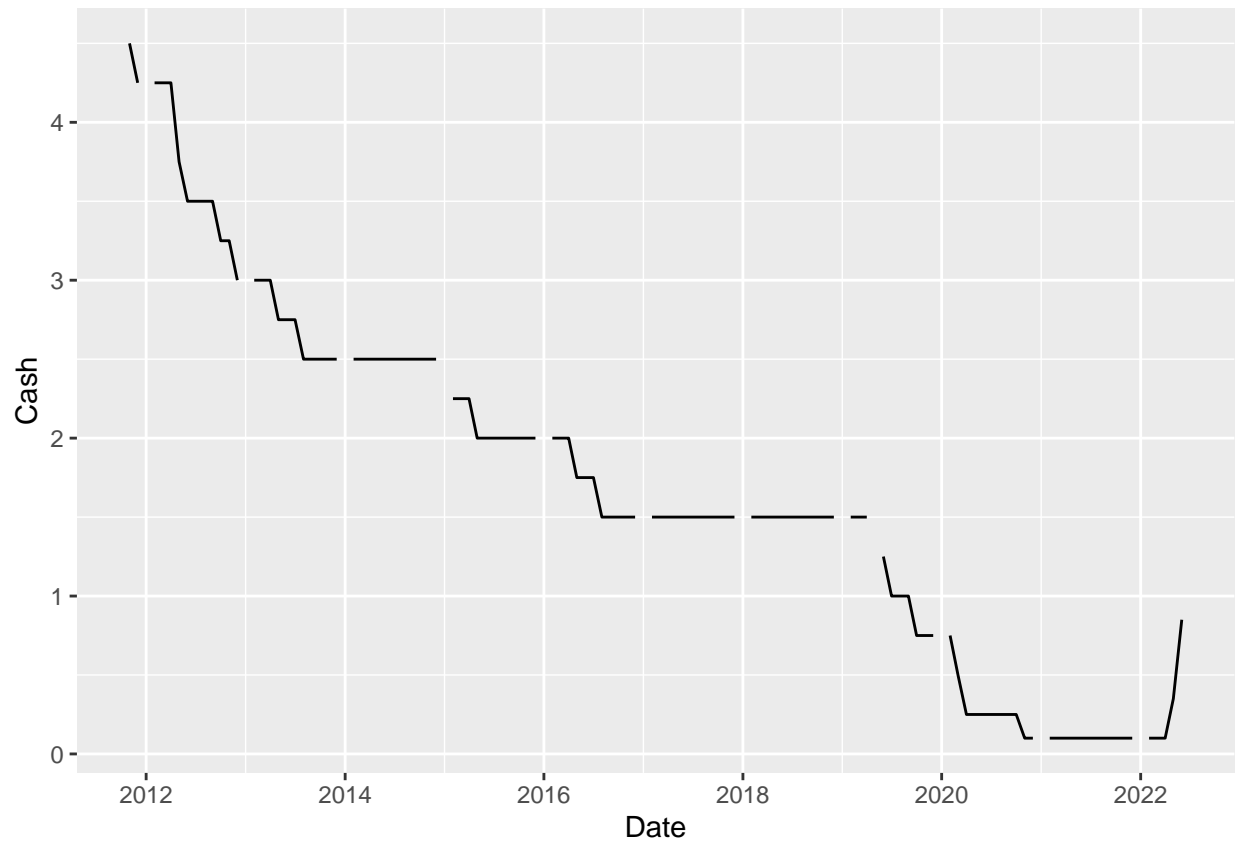
The following graph shows loan commitments has dropped from mid 2017 and the pattern continues till mid 2019. After that, the loan commitments again increased for some time and dropped till few months of 2020. After that, the loan commitments took a sharp incline.

```
ggplot(merge_dataset) +  
  geom_line(aes(x=Date, y=Total_loans))
```



The cash rate pattern seems quite opposite though. The government seems to have gradually reducing cash rate till 2022 but slightly increasing after 2022.

```
ggplot(merge_dataset) +  
  geom_line(aes(x=Date, y=Cash))
```



Unemployment rate seems to be highest in mid 2020. Can it be people are more free at that time to buy houses? See below

```
ggplot(merge_dataset) +  
  geom_line(aes(x=Date, y=Unemployment_rate))
```

