608SemesterProject

Charley Ferrari

Tuesday, April 14, 2015

First, lets read the data in and libraries.

```
library(ggplot2)
library(ggthemes)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following object is masked from 'package:stats':
##
##     filter
##
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
setwd("E:/Downloads/Courses/CUNY/SPS/Git/IS 608 Knowledge and Visual Analytics/Semester Project")

agent <- read.csv("agentdataprop.csv")

econ <- read.csv("econdataprop.csv")

econ$Date <- strptime(econ$Date, format="%m/%d/%Y")
```

I've chosen to use the company visit scores as my base data. On page 64 of the supporting documentation, seven of the eleven scores are mapped to macroeconomic concepts:

Demand and Output

Total Demand : Private Sector Nominal Output

Exports : Exports

Investment : Business Investment

Factor Utilisation

Employment : Private Sector Unemploymen

Costs and Prices

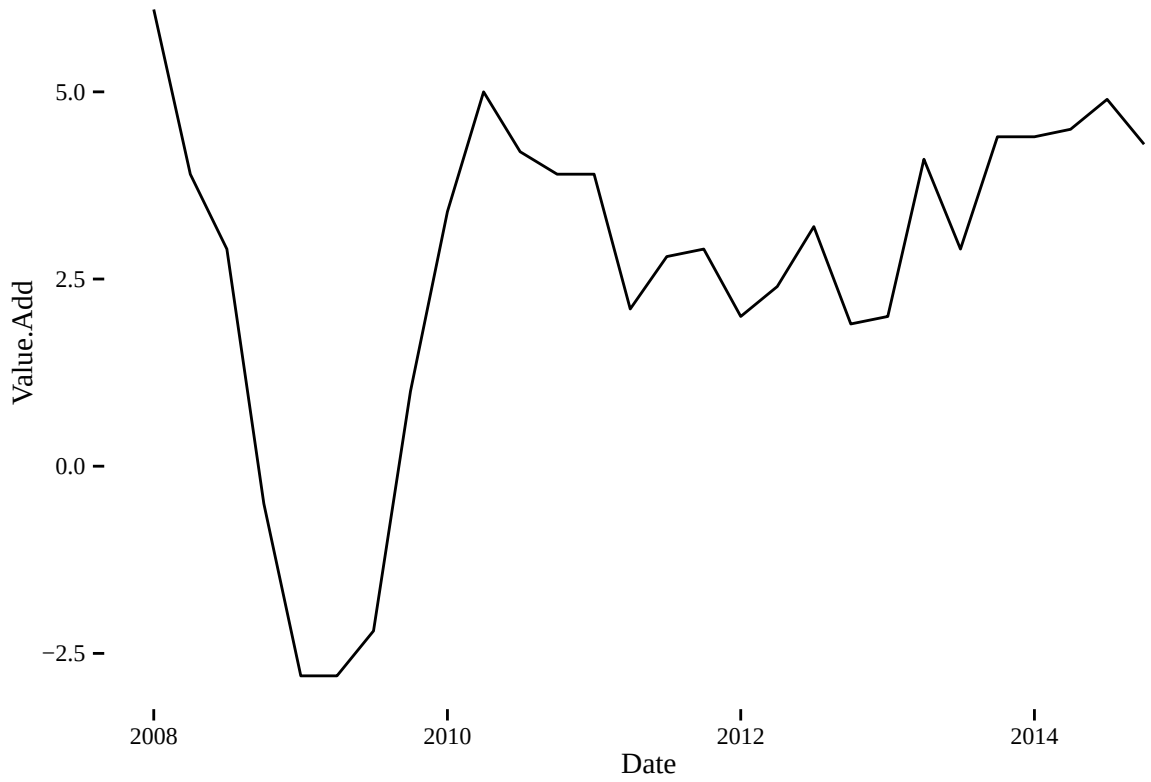
Total Labour Costs : Average Weekly Earnings Total Pay

Pay : Average Weekly Earnings Regular Pay

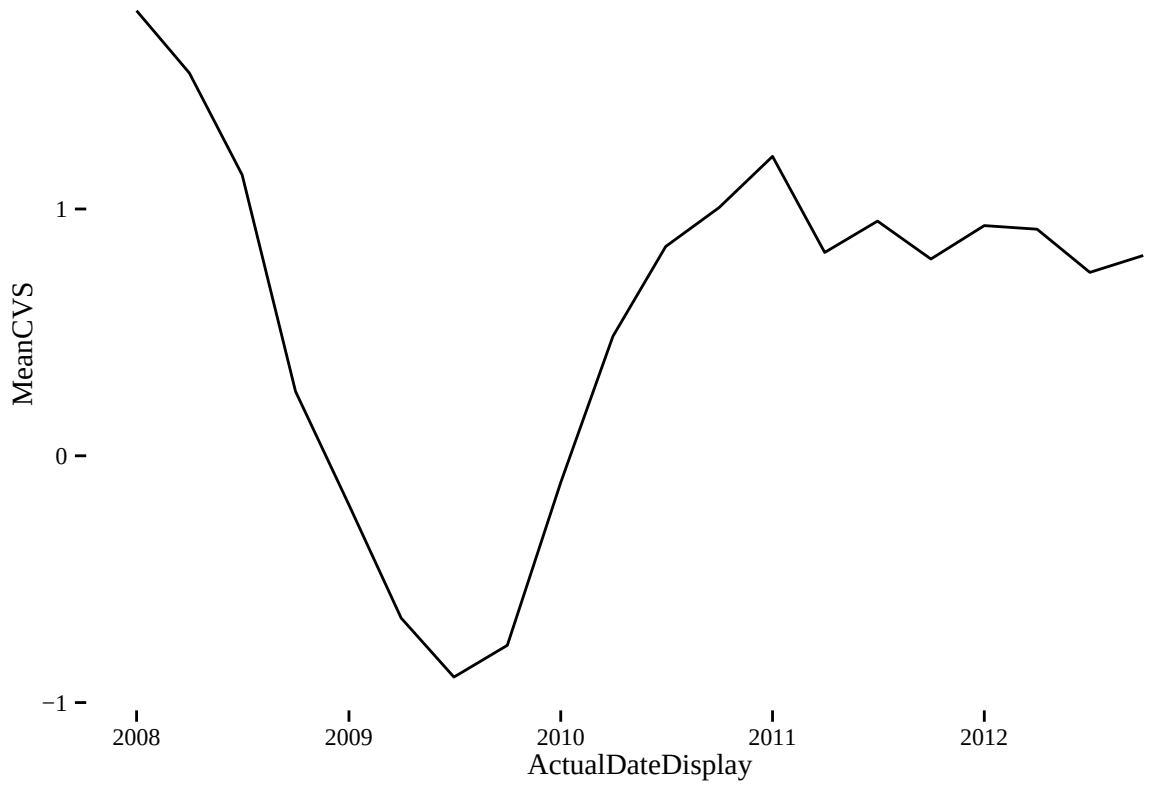
Profits : Gross Operating Surplus

My visualization will start out with two line graphs: On the top, a graph of the economic concept, and below, the aggregate company visit scores

```
agent01 <- agent %>%
  select(ActualDateDisplay, DemandScore) %>%
  group_by(ActualDateDisplay) %>%
  summarize(MeanCVS = mean(DemandScore))

agent01$ActualDateDisplay <- strptime(agent01$ActualDateDisplay, format="%m/%d/%Y")

agent01 <- agent01[order(agent01$ActualDateDisplay),]

ggplot(econ, aes(x=Date, y=Value.Add)) +
  geom_line() +
  theme_tufte()
```

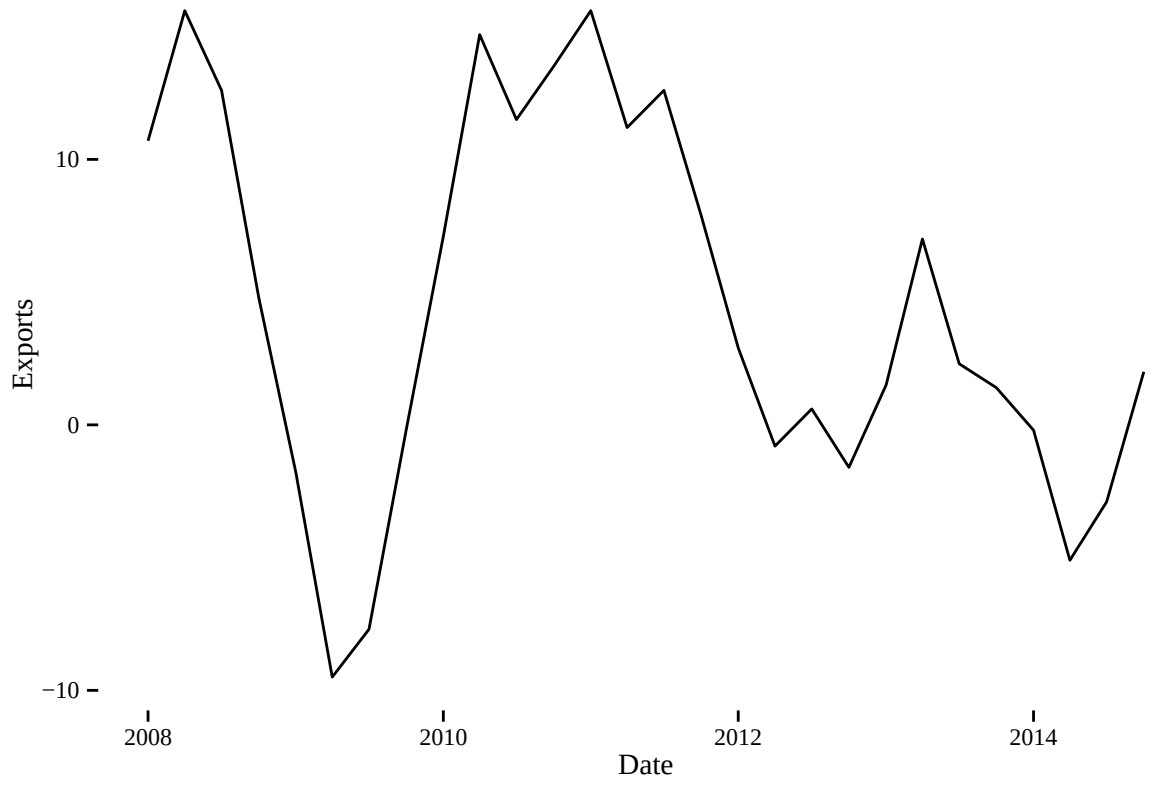



```
ggplot(agent01, aes(x=ActualDateDisplay, y=MeanCVS)) +
  geom_line() +
  theme_tufte()
```

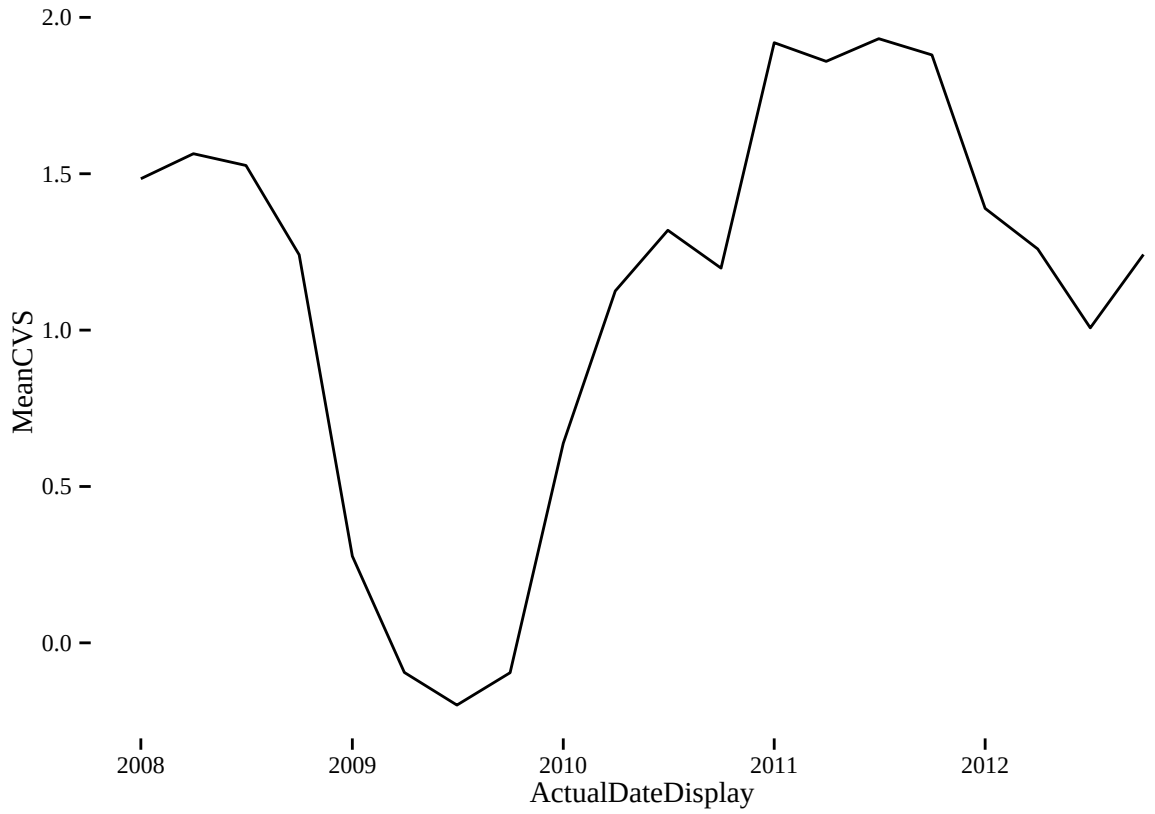


At this stage, users would be able to choose their concept from a drop down menu. If they wanted to select exports for example:

```
agent02 <- agent %>%
  select(ActualDateDisplay, ExportScore) %>%
  group_by(ActualDateDisplay) %>%
  summarize(MeanCVS = mean(ExportScore))

agent02$ActualDateDisplay <- strptime(agent02$ActualDateDisplay, format="%m/%d/%Y")

agent02 <- agent02[order(agent02$ActualDateDisplay),]

ggplot(econ, aes(x=Date, y=Exports)) +
  geom_line() +
  theme_tufte()
```



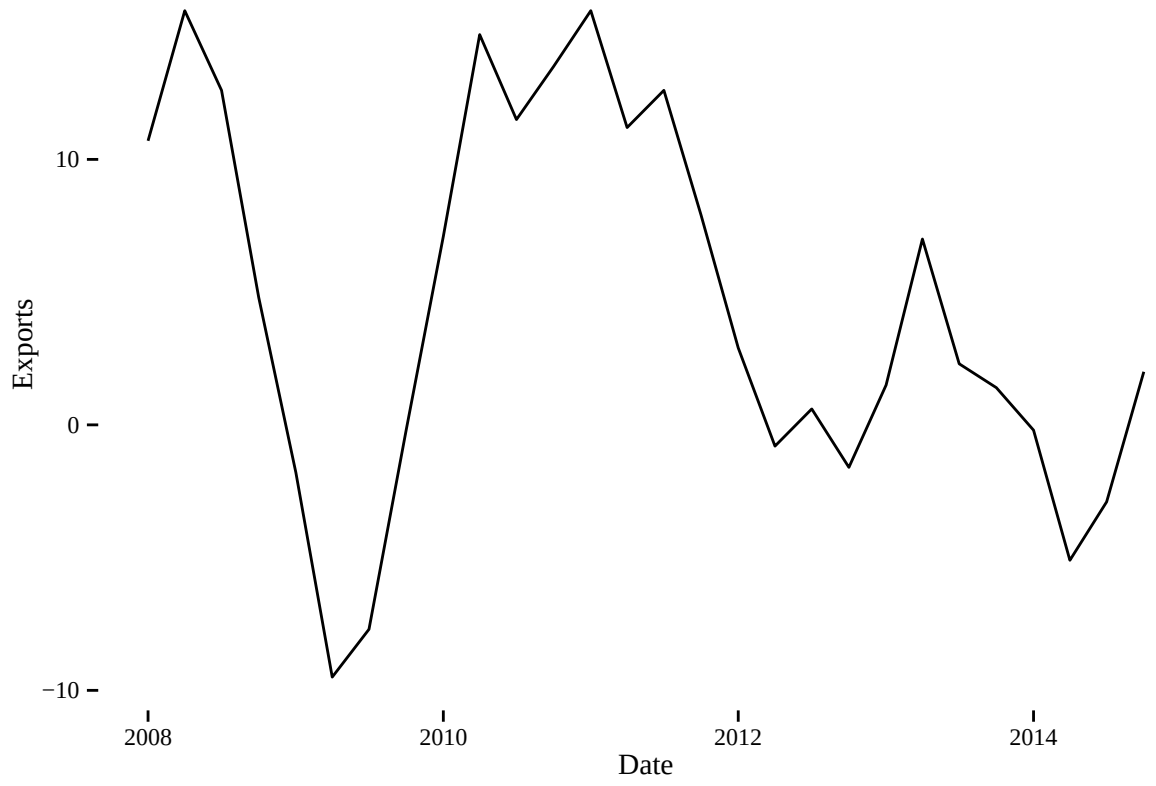

```
ggplot(agent02, aes(x=ActualDateDisplay, y=MeanCVS)) +
  geom_line() +
  theme_tufte()
```



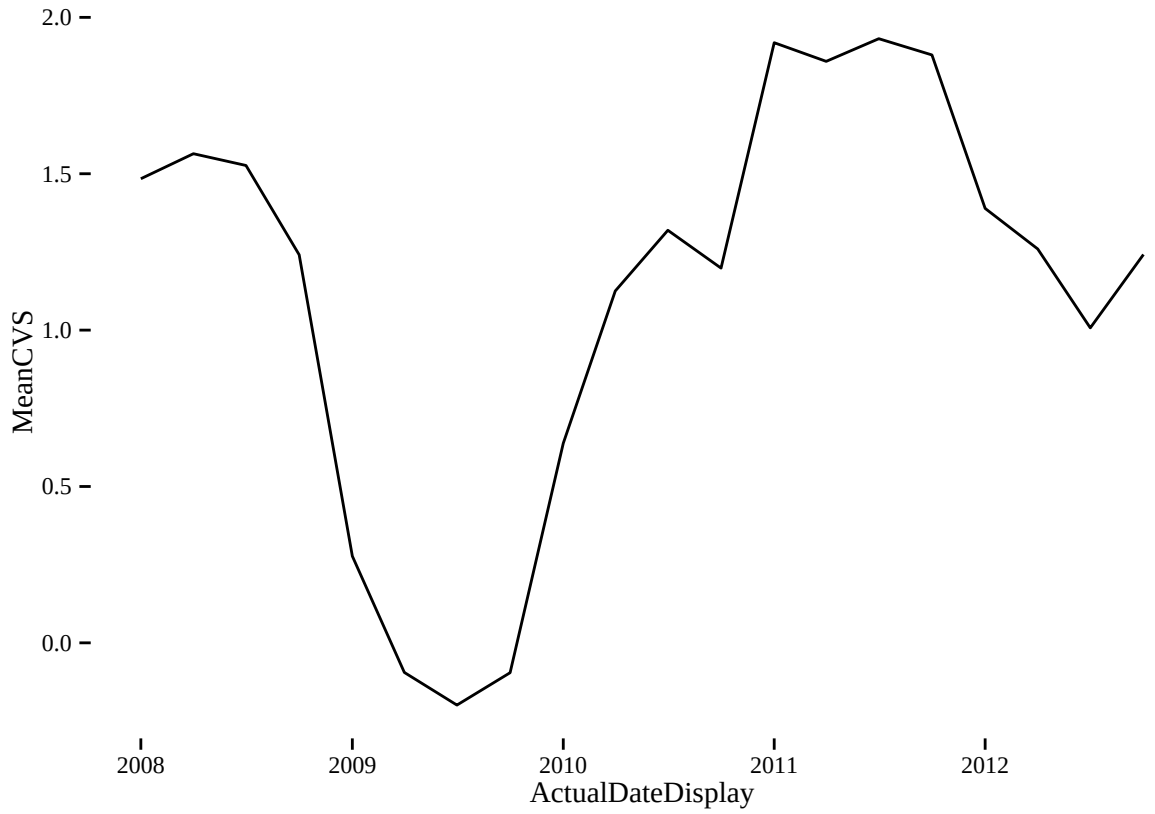
Instead of displaying the distribution of the CVS as a pop up, I think it would make more sense to display it as a third panel below these two graphs. I'd ideally like to have the graph show values when the user hovers over them, and allow the user to select a specifc time period by clicking on the point being hovered over.

So, lets say the user clicks on 01/01/2010:

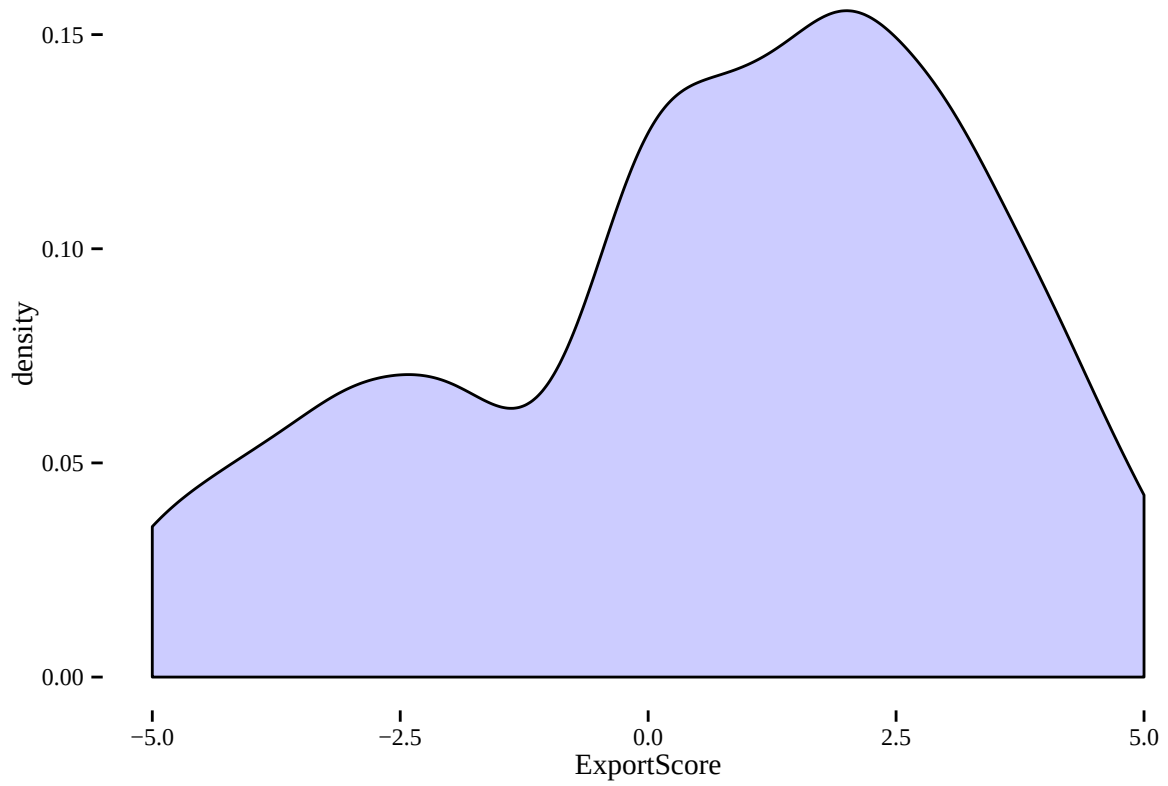
```
agent03 <- agent %>%
  filter(ActualDateDisplay == "1/1/2010") %>%
  select(Sector, ExportScore)

ggplot(econ, aes(x=Date, y=Exports)) +
  geom_line() +
  theme_tufte()
```




```
ggplot(agent02, aes(x=ActualDateDisplay, y=MeanCVS)) +
  geom_line() +
  theme_tufte()
```




```
ggplot(agent03, aes(x=ExportScore)) +
  geom_density(fill="blue", alpha=.2) +
  theme_tufte()
```

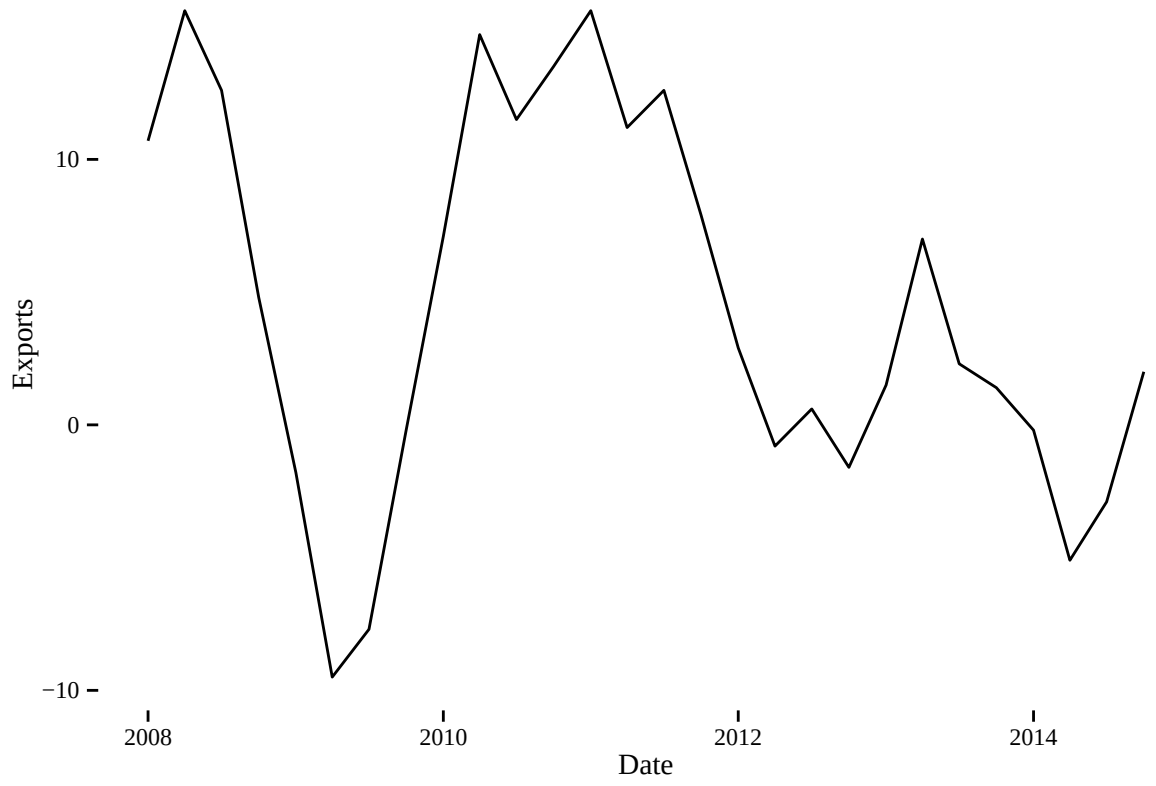


Now, in this third pane, I'd like to allow the user to select a sector via a drop down menu. Once the user selects a sector, the distribution will change, and the average CVS in pane 2 will change.

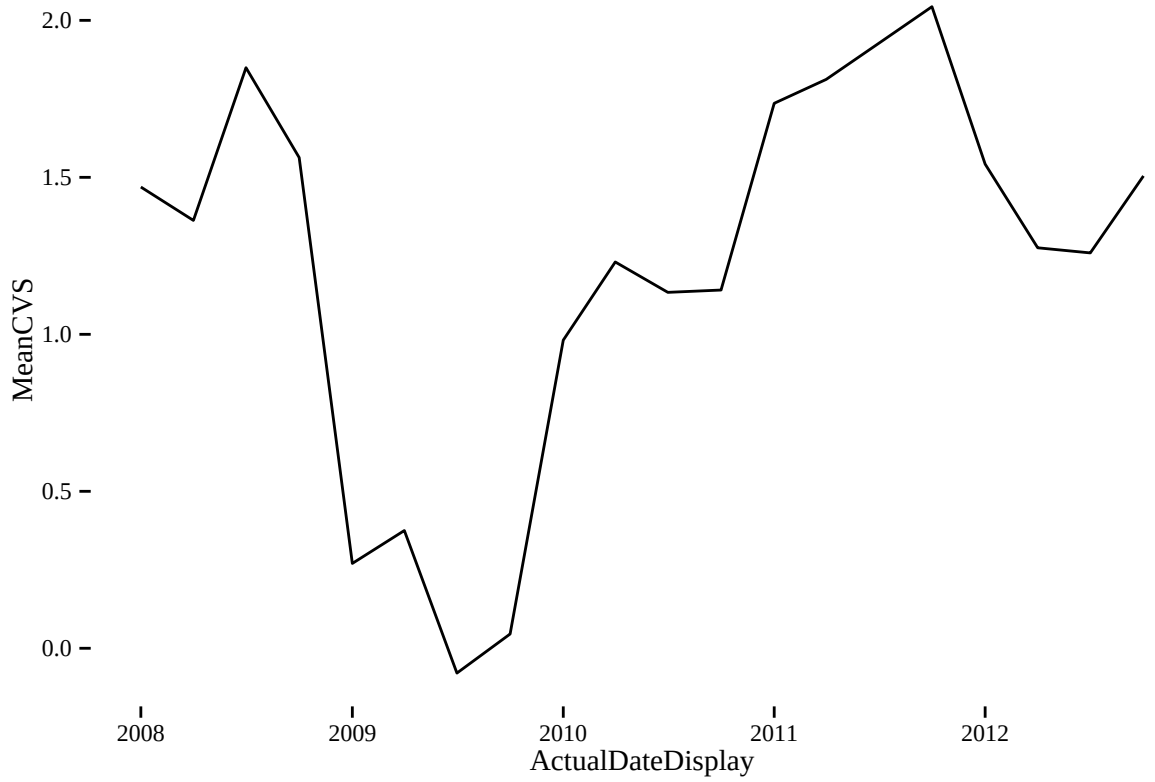
I'm still not sure how I will tweak the economic concepts. Some roll up nicely to the industries (such as employment), while others don't. Right now I'll just assume the economic concept remains the same, but I could just as easily imagine the selection of a sector will also change the headline economic concept

So, lets say the user chose "Business and Financial Services"

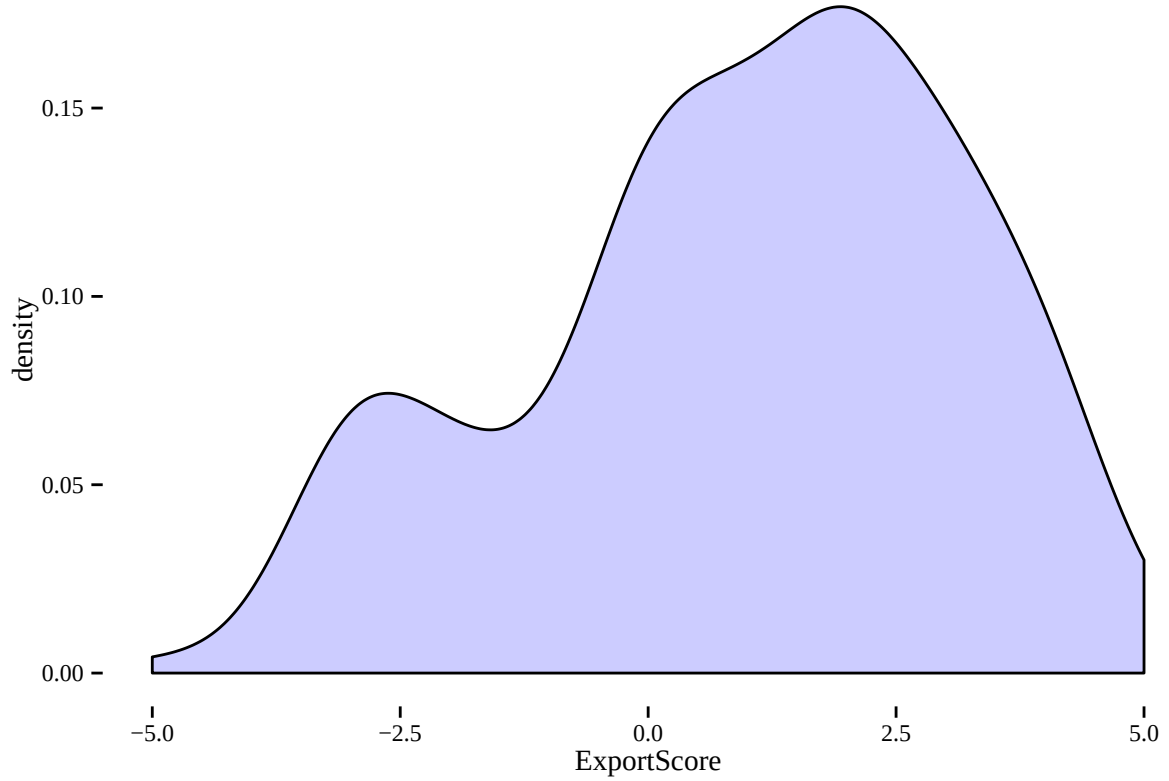
```
agent04 <- agent %>%
  filter(Sector == "Business and financial services") %>%
  select(ActualDateDisplay, ExportScore) %>%
  group_by(ActualDateDisplay) %>%
  summarise(MeanCVS = mean(ExportScore))

agent04$ActualDateDisplay <- strptime(agent04$ActualDateDisplay, format="%m/%d/%Y")

agent04 <- agent04[order(agent04$ActualDateDisplay),]

agent05 <- agent03 %>%
  filter(Sector == "Business and financial services")

ggplot(econ, aes(x=Date, y=Exports)) +
  geom_line() +
  theme_tufte()
```




```
ggplot(agent04, aes(x=ActualDateDisplay, y=MeanCVS)) +
  geom_line() +
  theme_tufte()
```




```
ggplot(agent05, aes(x=ExportScore)) +
  geom_density(fill="blue", alpha=.2) +
  theme_tufte()
```



There are a few other selections I could allow the user to make. Future scores are given in addition to backward looking scores, so that could just as easily be added in a drop down menu that affects all panes.

In reading through the documentation, I had another insight into how this data can be viewed: the distributions can be animated. Chart 4 shows the distribution of CVS profit scores in 2009 versus 2012. I think this could be viewed as an animation, showing how the density curve changes through time.

More interestingly, charts 5 and 8 show scatter plots just begging to be changed into Hans Rosling style animated charts.

In pane 3, I'd like to add a "play" button, that would allow the distribution to be animated, and return to the date chosen in pane 1 or 2. I'd also like to add a 4th pane that would allow the user to pick another concept to compare. This would allow the sector picked in pane 3 to apply to the scatter plot.