The ggplot() Command

- The command ggplot() initializes a ggplot object.
- It can be used to declare the input data frame for a graphic.
- It can also be used to specify the set of plot aesthetics intended to be common throughout all subsequent layers (unless specifically overridden).
- The actual plots are built with subsequent commands.
- ggplot() is typically used to construct a plot incrementally, using the + operator to add layers to the existing ggplot object.
- This is advantageous in that the code is explicit about which layers are added and the order in which they are added.
- We can define common aesthetics using the aes argument for this command.

Aesthetics

- Aesthetics are attributes that can be perceived on the graphic.
- Size, shape and colour are all examples of aesthetics.
- Each aesthetic can be mapped to a variable, or set to a constant value.

Geoms

Points, lines and bars are all examples of *geoms* (geometric objects)

Geom	Named Plot	
point	Scatter plot	
point	Bubble chart	(relationship of size to variable)
bar	Bar chart	
boxplot	Box-whisker plot	
line	Line chart	

Geom smooth

- Basic scatterplot seen previously. Add a smooth geom to the plot
- The "smooth" layer fits a smooth line through the data.
- The stat fits the data to a loess smoother, with semi-transparent ribbon for representing standard error.
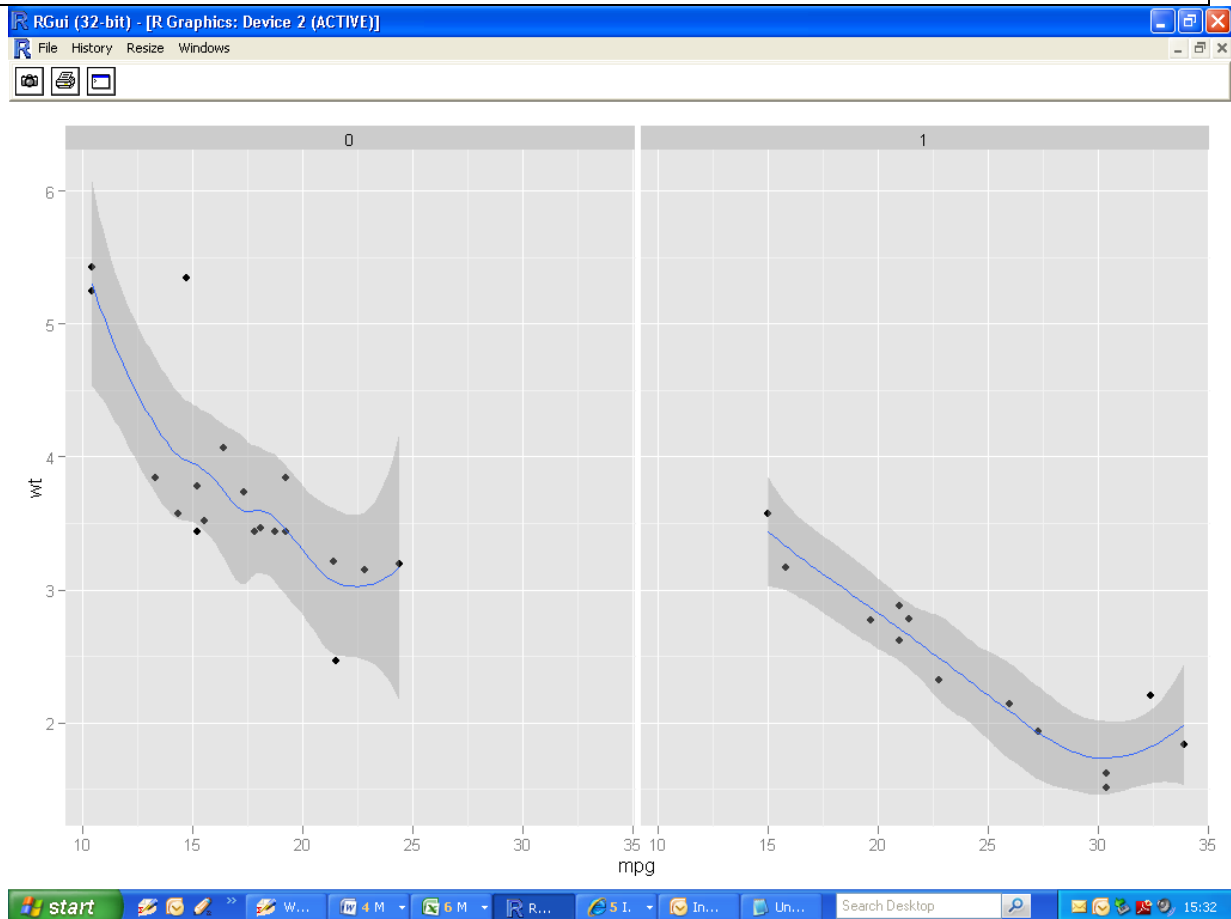
```
>
>qplot(mpg,wt,data=mtcars)+geom_smooth()
>
```
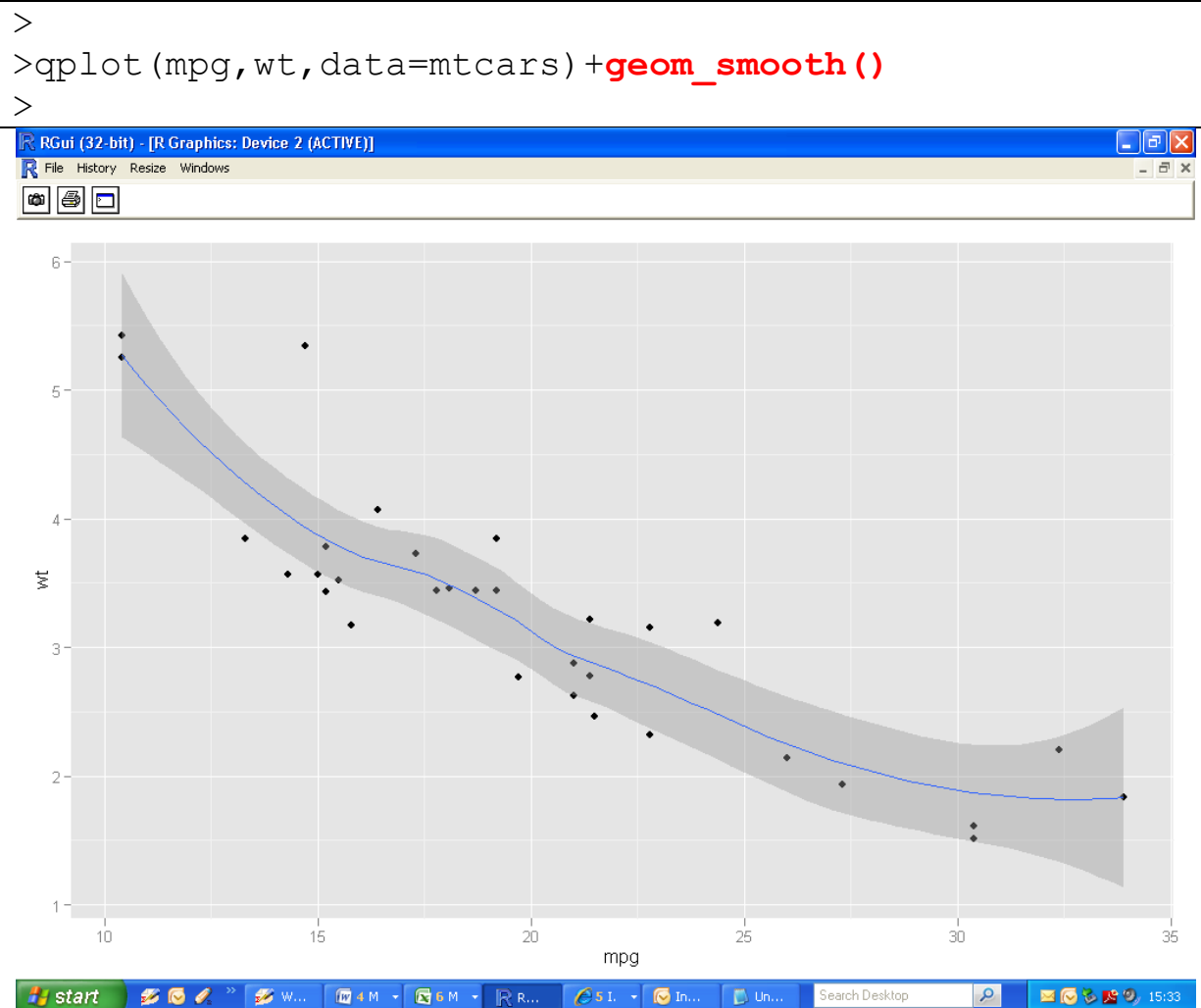

Warning Message

```
geom_smooth: method="auto" and size of largest group
is <1000, so using loess. Use 'method = x' to change
the smoothing method.
```

Additional Example (using faceting)

```
>
>qplot(mpg,wt,data=mtcars, facets = .~am)
+geom_smooth()
>
```
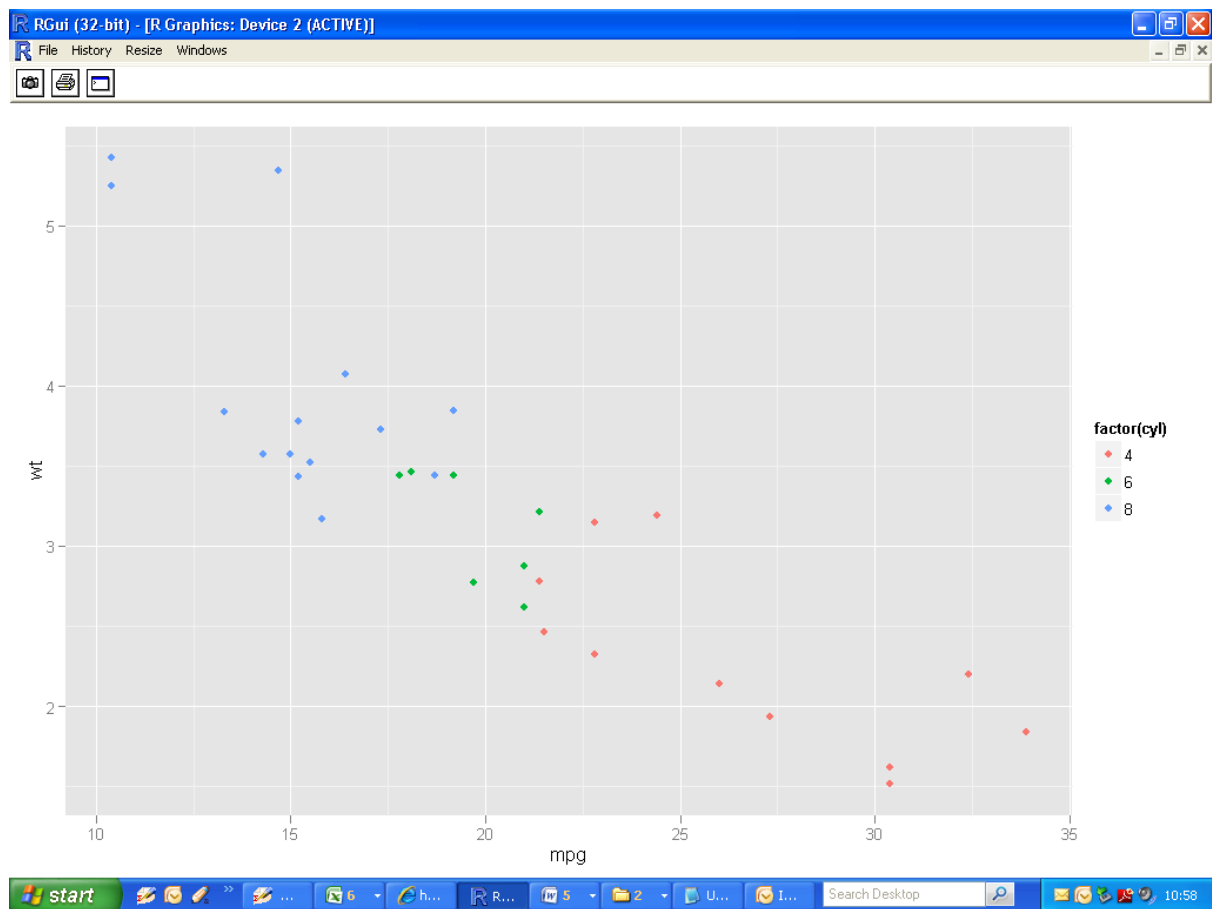
Using ggplot()

- Start off in similar manner to `qplot()`.
- The function `ggplot()` only creates a data object. There is no graphic yet. (Try out summary() on the object).
- Add layers to the data object to built up plot.

```
>P = ggplot(mtcars, aes(mpg, wt, colour=factor(cyl)))
```
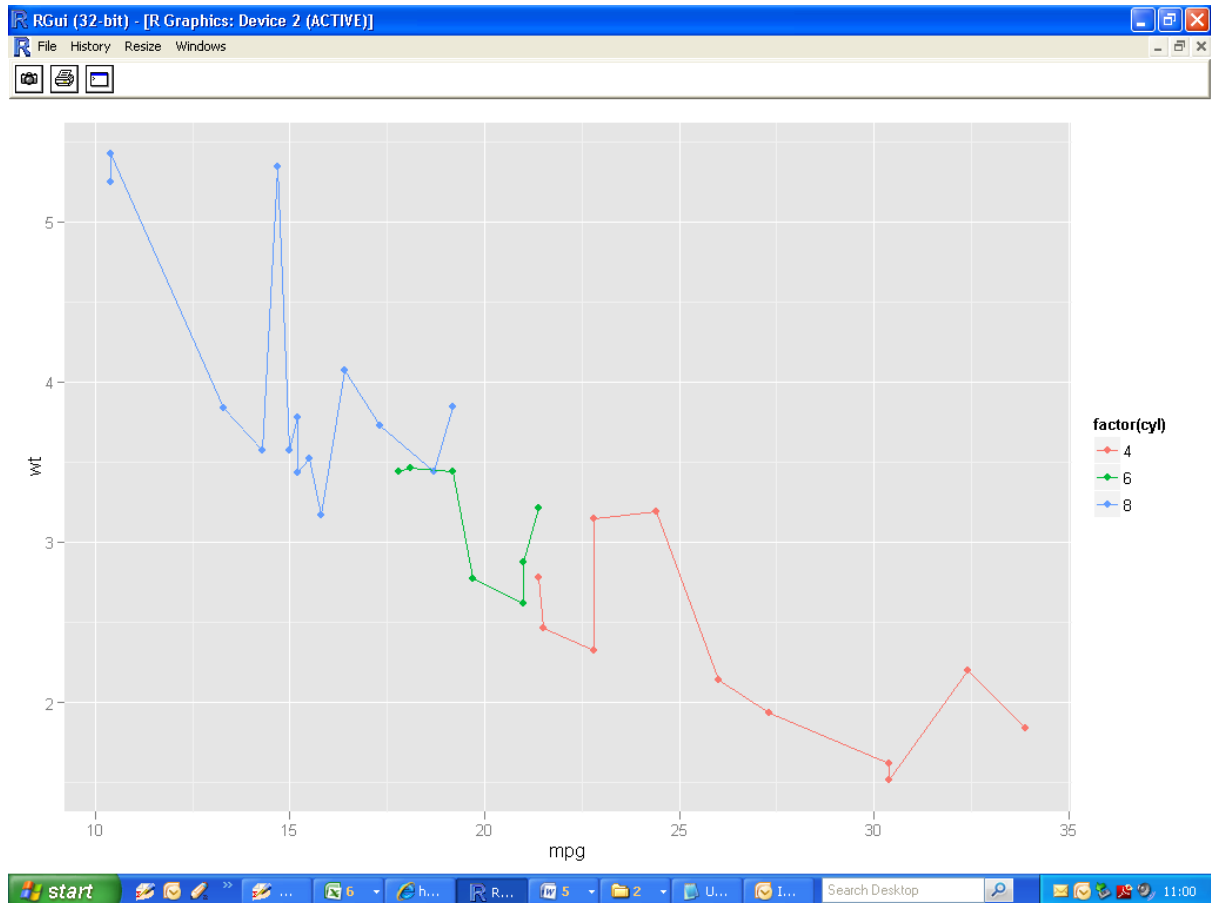
Add the first layer

- Both of the following commands are equivalent.

```
>P + geom_point()
>P + layer(geom="point")
```
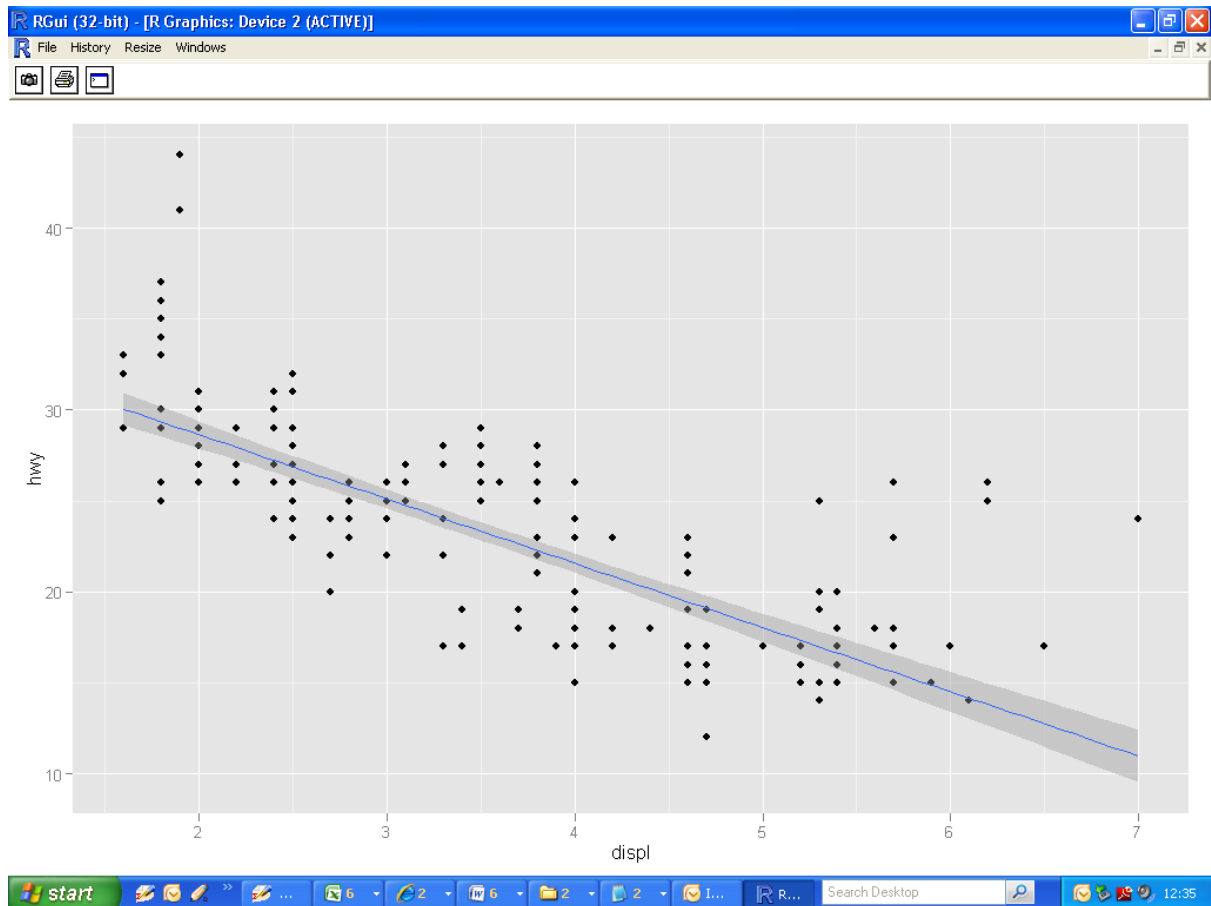
Add another layer

```
>P + geom_point()+ geom_line()
>P + layer(geom="point") + layer(geom="line")
```
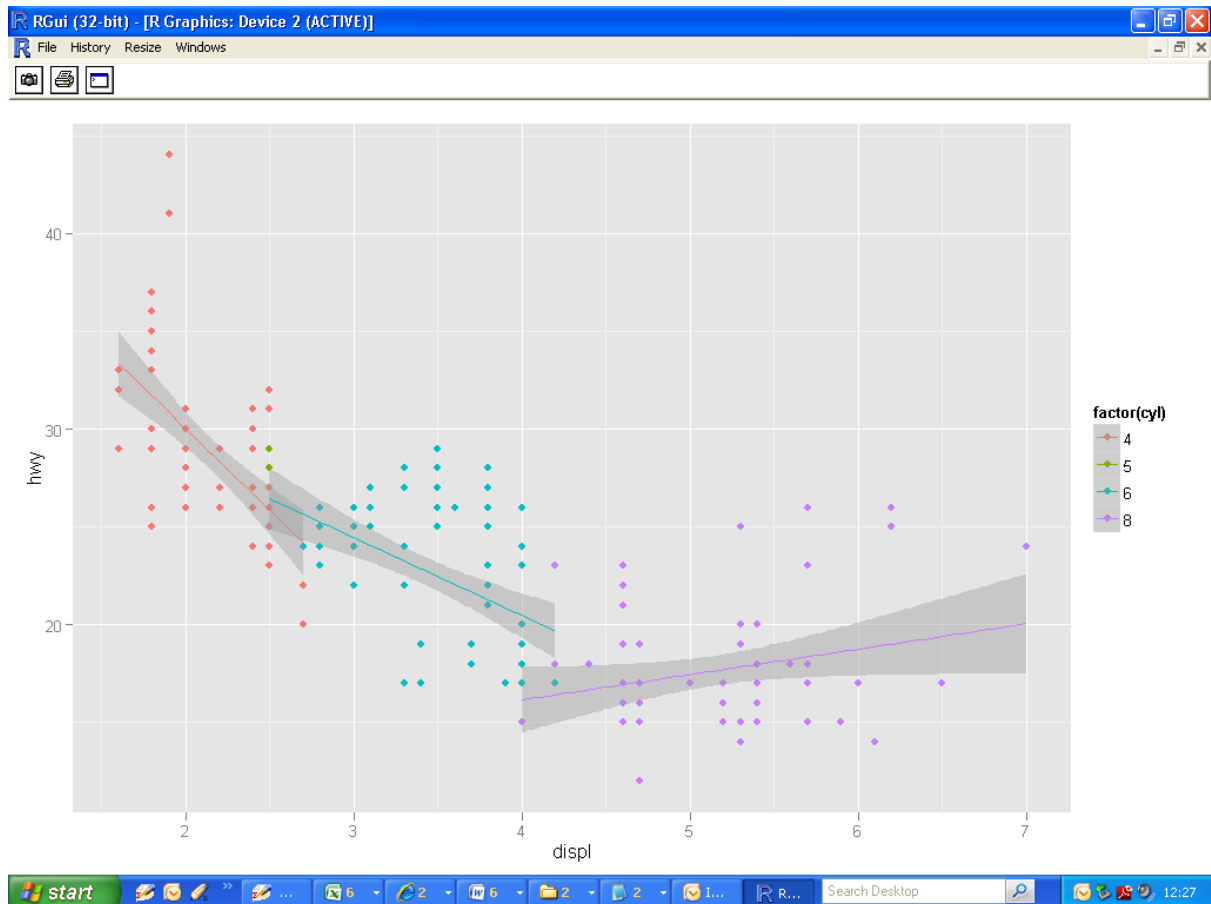
Linear Regression (entire set)

```
> ggplot(mpg,aes(displ,hwy))
+ geom_point()
+ geom_smooth(method="lm")
```

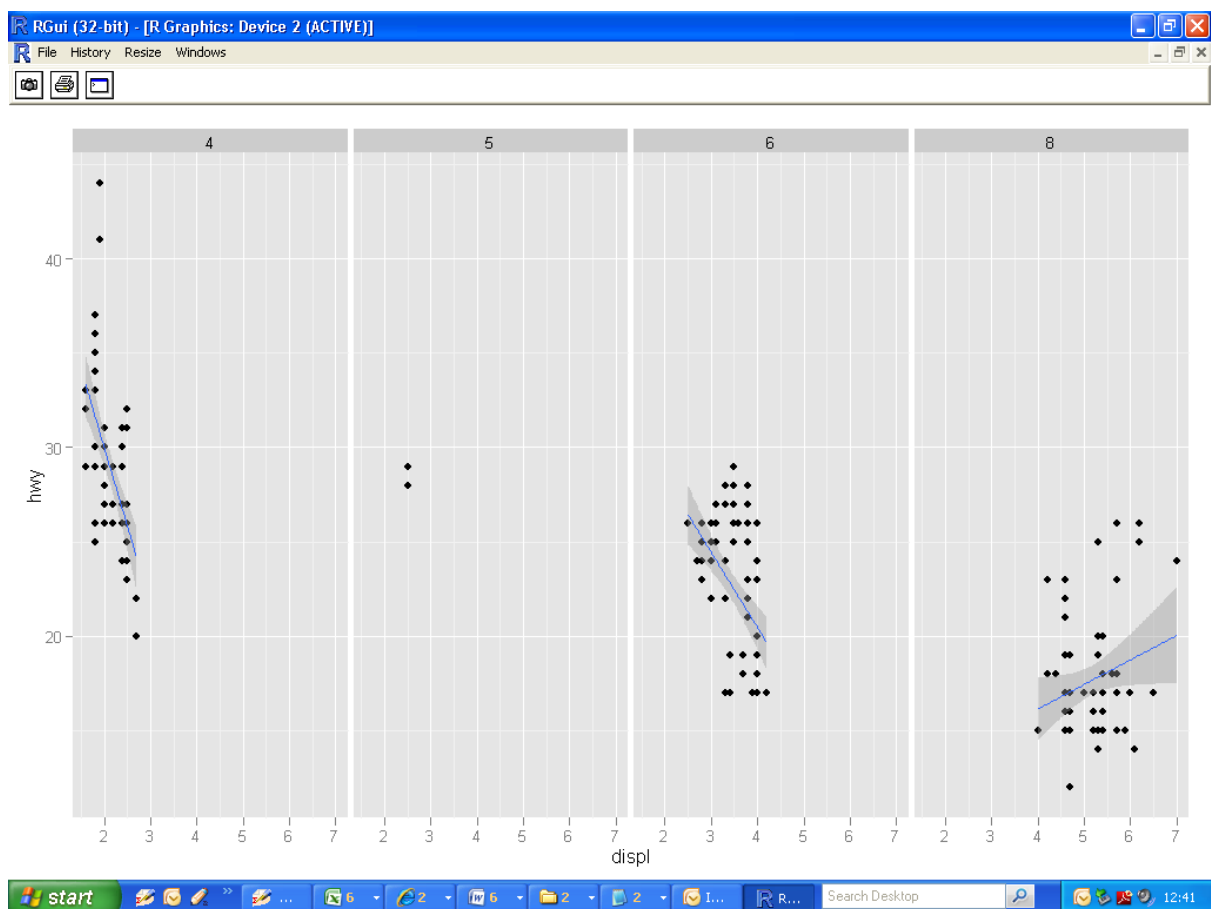## Linear Regression (using grouping)

```
> ggplot(mpg,aes(displ,hwy,colour=factor(cyl)))
+ geom_point()
+ stat_smooth(method="lm")
```

Linear Regression (using faceting)

```
> P = ggplot(mpg,aes(displ,hwy))+ geom_point()
+ stat_smooth(method="lm")
> P+ facet_grid(.~cyl)
```
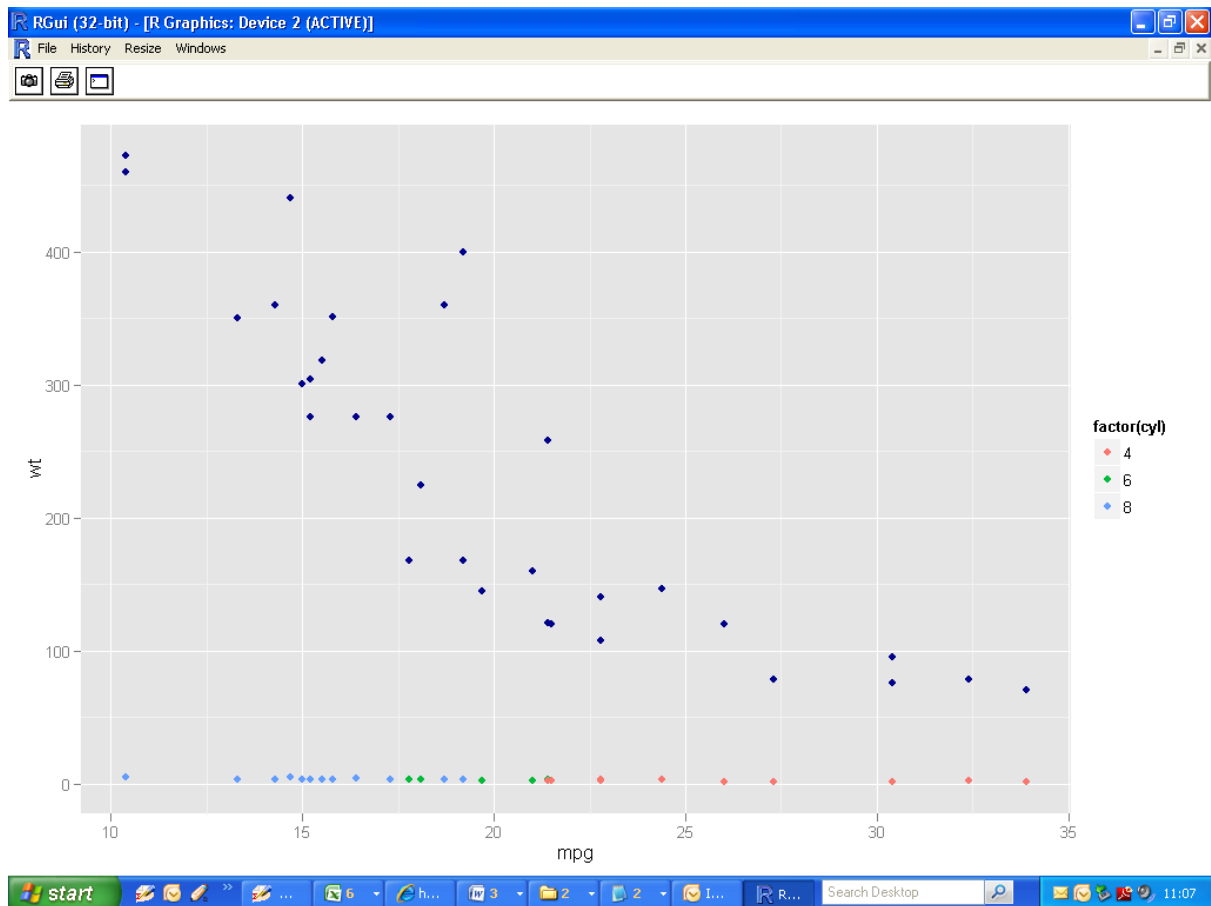
- `facet_grid(cyl ~ .)` :More or less the same graph, but on top of each other.
- Can adjust scales!

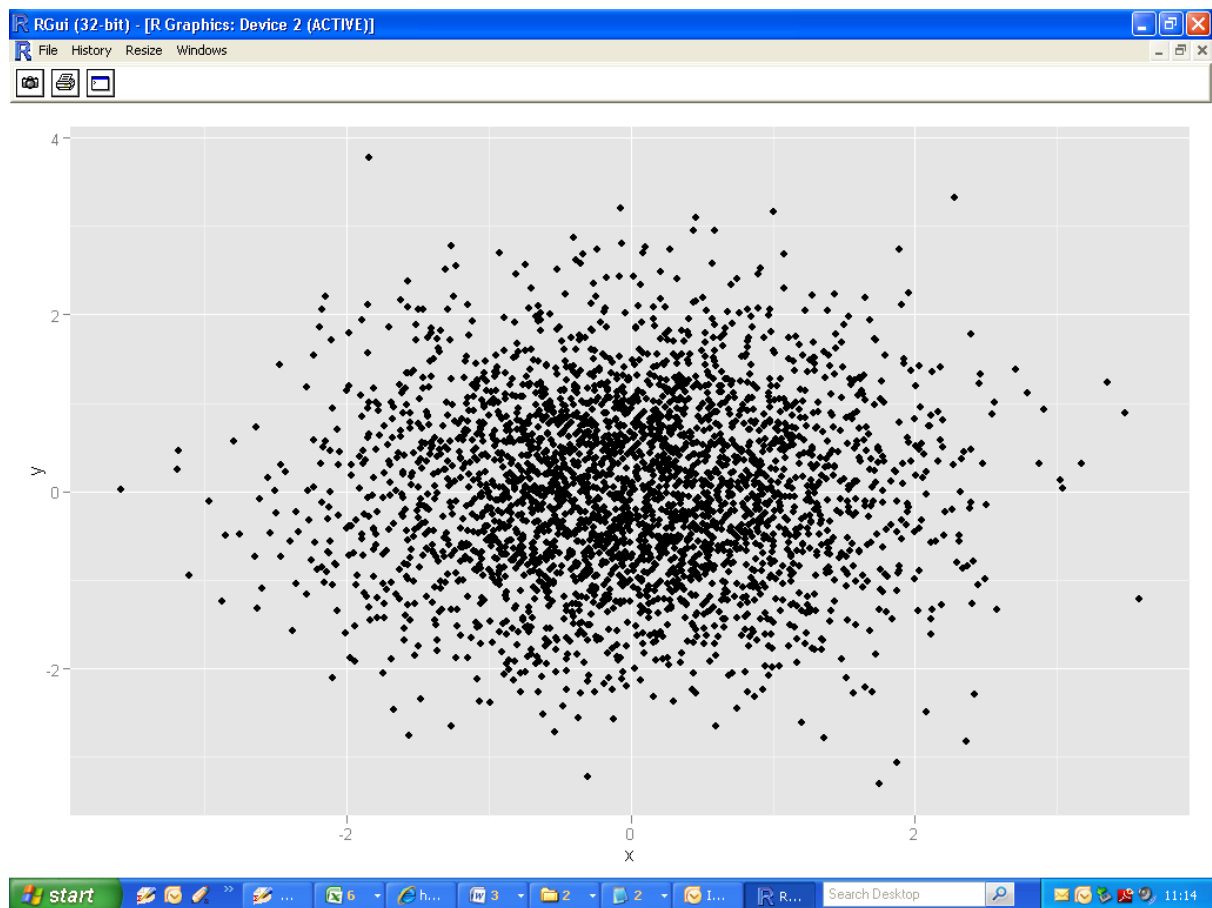## Add an additional layer with different mapping

```
P + geom_point() +
geom_point(aes(y=disp),colour="darkblue")
```

- Strange plot: Faceting is useful for situations like this

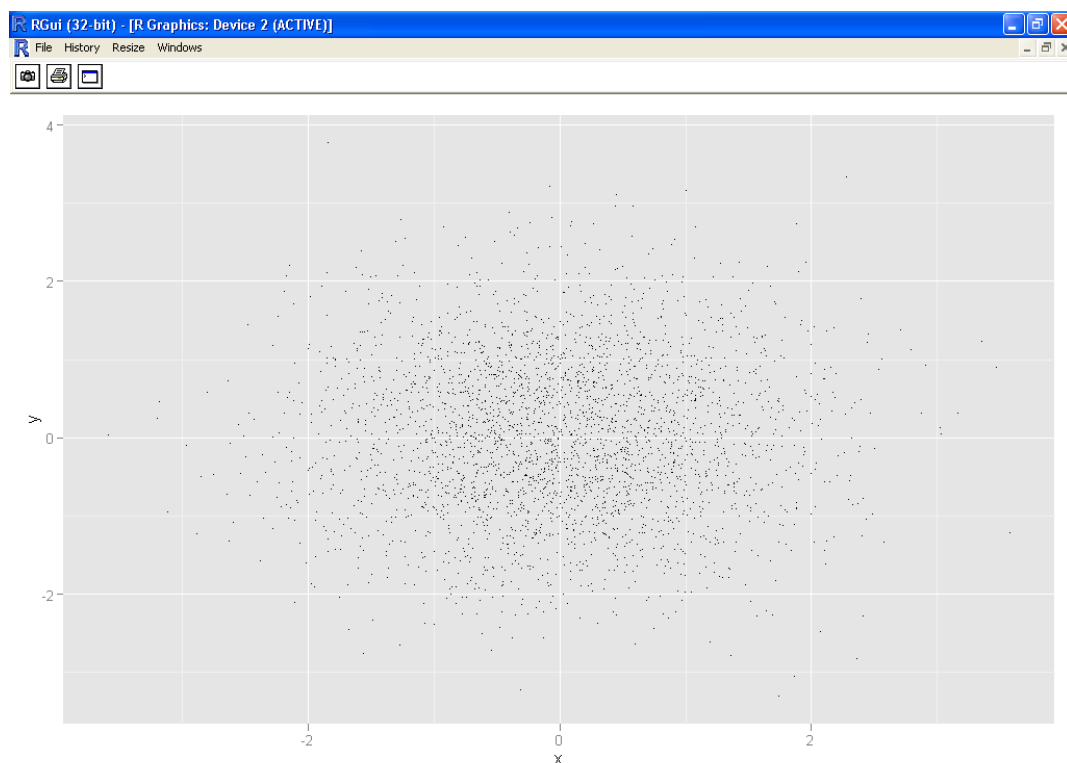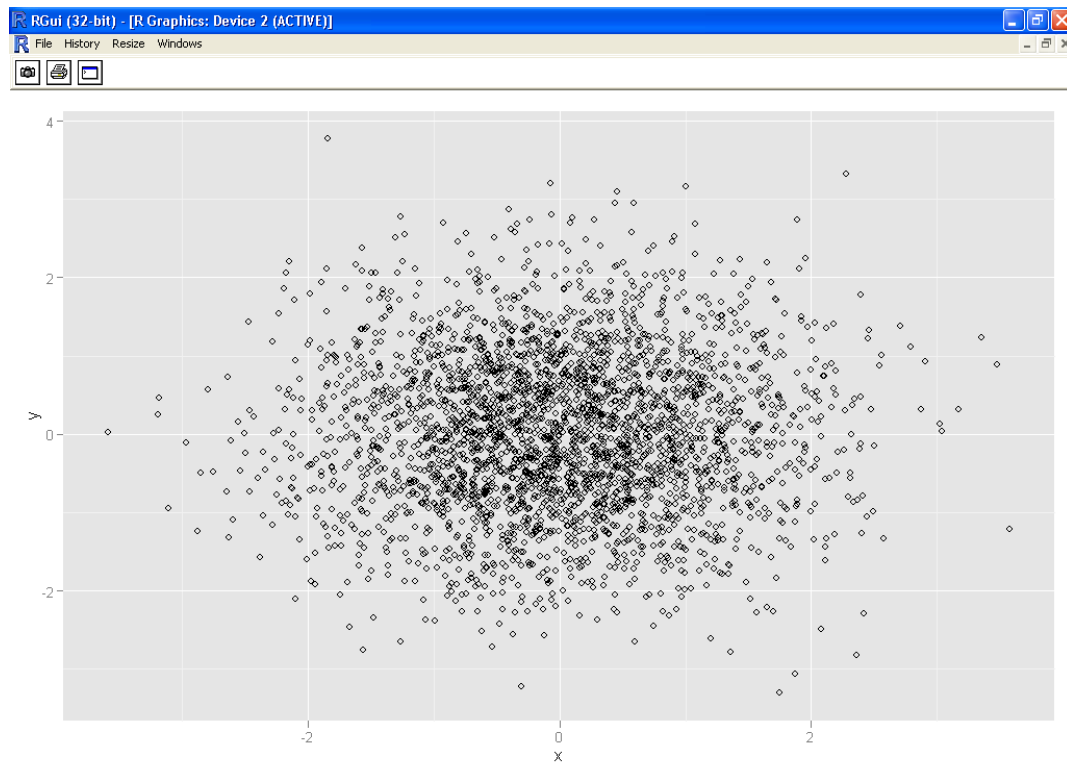## Dealing with overplotting

```
>DFnorm<- data.frame(x=rnorm(3000), y=rnorm(32000))
>p.norm <- ggplot(DFnorm, aes(x,y))
>
>p.norm + geom_point()
```



```
p.norm + geom_point(shape = 1)
p.norm + geom_point(shape = ".")
```

Hollow Glyphs / Narrow Glyphs

The

Transparency (alpha blending)

```
>p.norm + geom_point(alpha=(0.25),colour="blue")
>
```
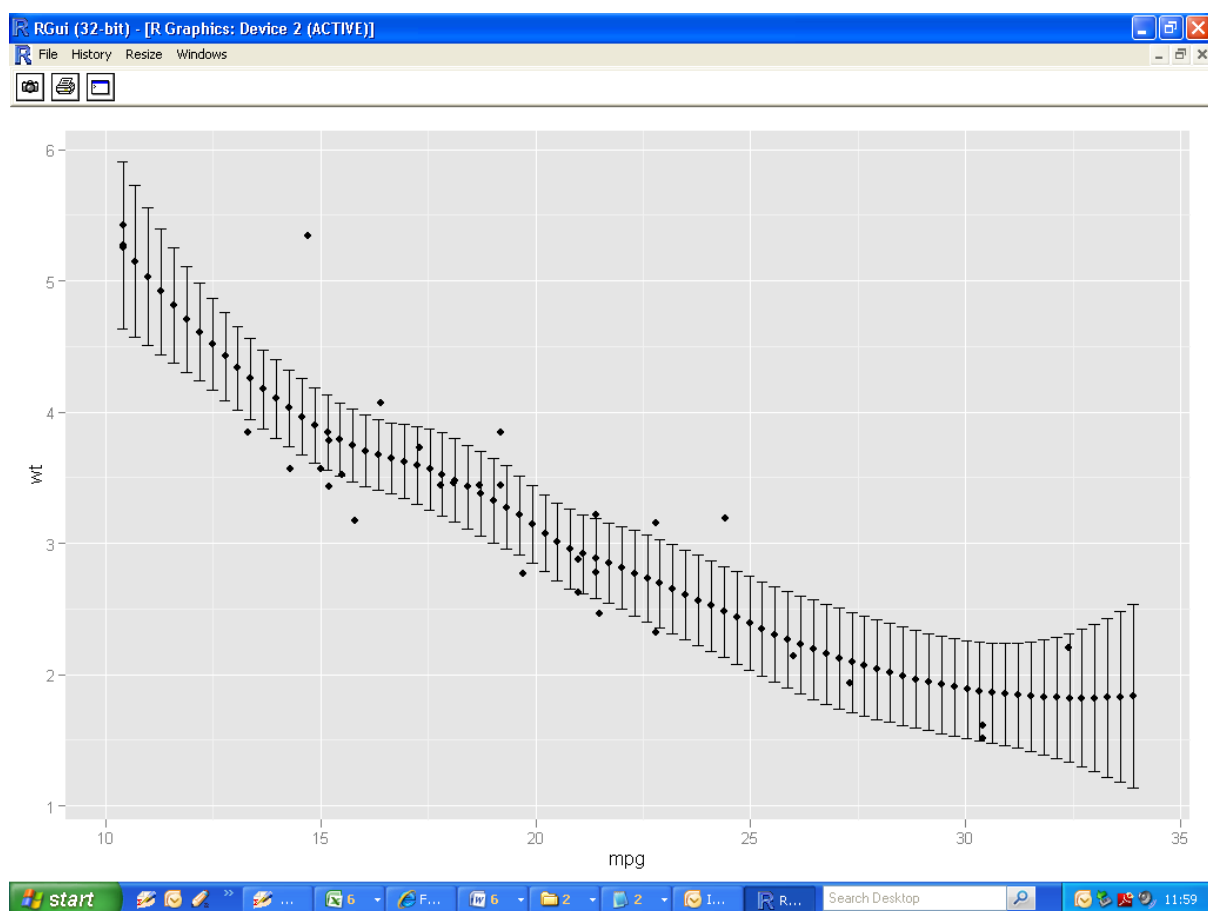
Adding statistics to your plot

```
>
>qplot(mpg,wt,data=mtcars)+stat_smooth()
>
```

- Same plot as before.
- `stat_smooth()` equivalent to `geom_smooth()` and `geom_ribbon()` by default
- Use different geoms: Points and Error bars(probably bad idea)

```
>
>qplot(mpg,wt,data=mtcars)+
stat_smooth(geom="point") +
stat_smooth(geom="errorbar")
>
```

Geoms and Stats

- Geoms always have default statistics associated with them.
- Stats always have default geoms associated with them.
- Many are interchangeable (e.g. smooth and boxplot)

Boxplots

```
>ggplot(mpg,aes(class,hwy))+ stat_boxplot()
>ggplot(mpg,aes(class,hwy))+ geom_boxplot()
>
```