

QTM 151

Week 8 - qplot

Umberto Mignozzetti

Recap

We learned:

- programming methods
- `dplyr *_join` methods: joining data
- `tidyr` methods: reshape datasets
- `forcats` methods: working with categorical variables
- `lubridate` methods: to work with dates

Great job!!

Do you have any questions about any of these contents?

Today we are going to talk about **qplot** (package for quick plots in ggplot2)

Our GitHub page is: <https://github.com/umbertomig/qtm151>

qplot

qplot

- `qplot`, as the name says, stands for quick plots.
- It is great to generate `ggplot` graphs in a hurry.
- And the graphs are indeed `ggplot` objects: you can still add layers to it!

qplot - Geoms Available

- Which graphs can we generate?
 - **"point"**: scatterplots.
 - **"line"**: line plot.
 - **"histogram"**
 - **"boxplot"**
 - **"density"**
 - **"bar"**: barplot.
 - **"smooth"**: Fits a smooth line.
 - **"dotplot"**: dotplot.

qplot - Options for Customization

- And there are plenty of quick options to customize the graphs.
 - `data`: Specify the data-frame.
 - `main`: Title.
 - `xlab`, `ylab`: x and y axis labels.
 - `color`: Controls the color of the lines/points.
 - `fill`: Controls the color of areas (e.g. for histograms).
 - `size`: Controls the size of points.
 - `shape`: The shape of points ("circle", "square", "triangle", etc...)
 - `alpha`: Controls the level of transparency of points/lines/fills.
 - `lwd`: Line width.
 - `lty`: Line type ("solid", "dashed", "dotted", etc...).
 - `facets`: Split up the data into multiple plots.

Loading tidyverse

```
# Load tidyverse
```

```
library(tidyverse)
```

```
## — Attaching packages ————— tidyv
```

```
## ✓ ggplot2 3.3.5      ✓ purrr 0.3.4
```

```
## ✓ tibble 3.1.6       ✓ dplyr 1.0.8
```

```
## ✓ tidyr 1.2.0        ✓ stringr 1.4.0
```

```
## ✓ readr 2.1.2        ✓ forcats 0.5.1
```

```
## Warning: package 'tidyr' was built under R version 4.1.2
```

```
## Warning: package 'readr' was built under R version 4.1.2
```

```
## Warning: package 'dplyr' was built under R version 4.1.2
```

```
## — Conflicts ————— tidyverse_c
```

```
## x dplyr::filter() masks stats::filter()
```

Loading datasets

```
# Loading tips dataset
```

```
tips ← read.csv('https://raw.githubusercontent.com/umbertomig/qtn  
head(tips, 2)
```

```
##      obs totbill  tip sex smoker day  time size  
## 1     1   16.99 1.01  F      No Sun Night    2  
## 2     2   10.34 1.66  M      No Sun Night    3
```

```
# Loading PErisk dataset
```

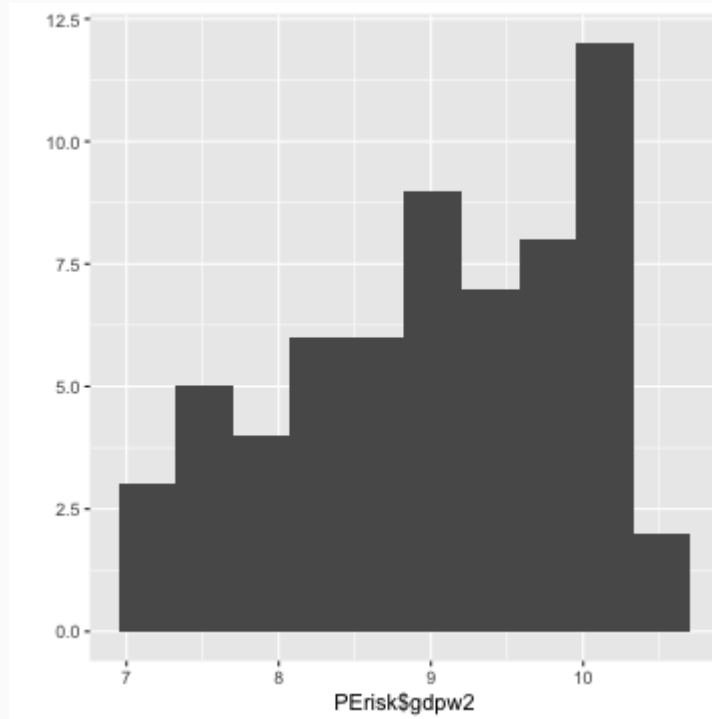
```
PErisk ← read.csv('https://raw.githubusercontent.com/umbertomig/c  
head(PErisk, 2)
```

```
##      country courts      barb2 prsexp2 prscorr2      gdpw2  
## 1 Argentina      0 -0.7207754      1      3  9.69017  
## 2 Australia      1 -6.9077550      5      4 10.30484
```


Plots for Continuous Variables

qplot - Histograms

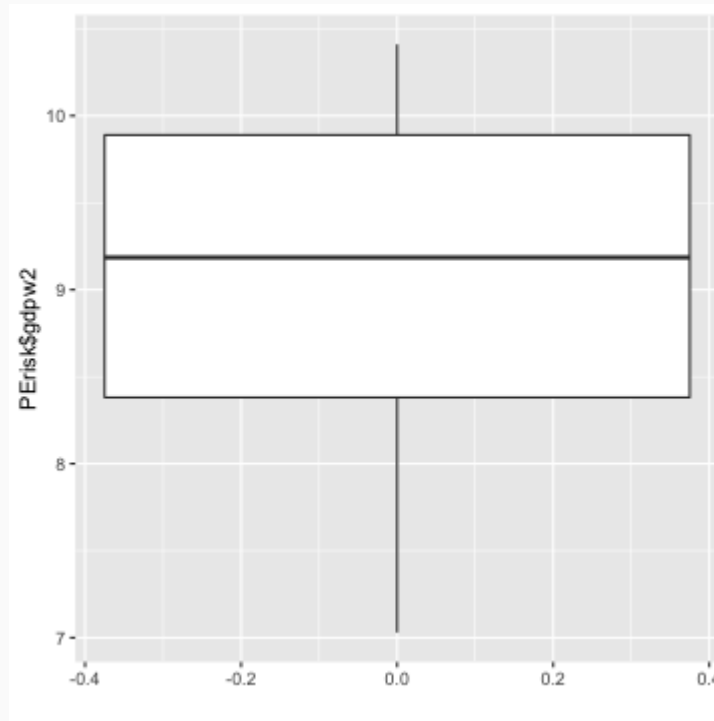
```
qplot(PErisk$gdpw2, geom = "histogram", bins = 10)
```



- **Your turn:** make a histogram of the `tip` variable in the `tips` dataset.

qplot - Box-plots

```
# Box-plot of log of per capita gdp  
qplot(y = PErisk$gdwp2, geom = "boxplot")
```



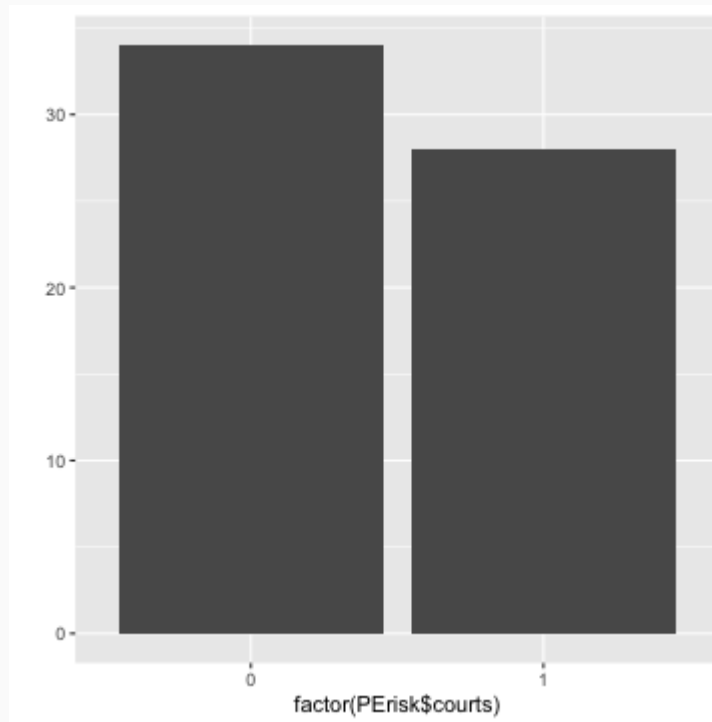
- **Your turn:** make a boxplot of the `totbill` variable in the `tips` dataset.

Plot for Discrete Variables

qplot - Bar-Plots

```
# Bar-plot of courts
```

```
qplot(factor(PERisk$courts), geom = "bar")
```



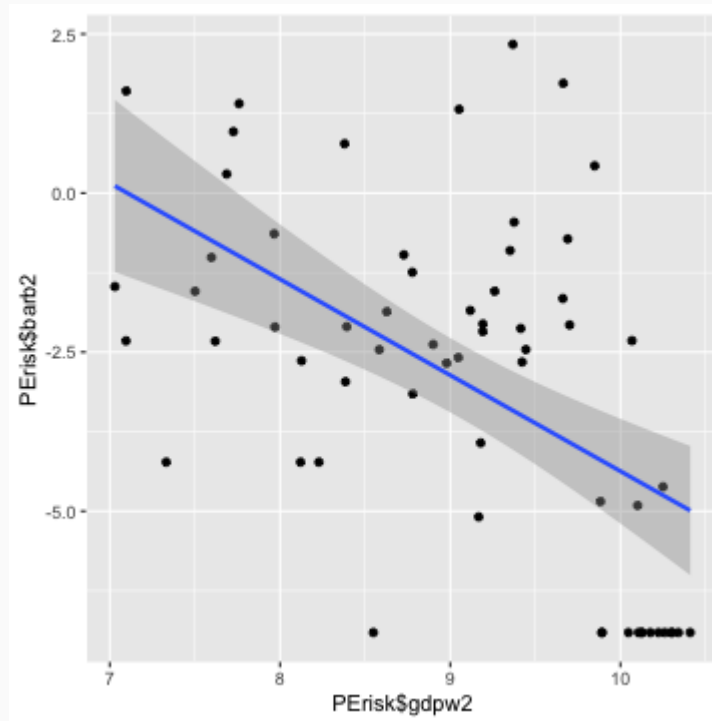
- **Your turn:** make a bar-plot of the `smoker` variable in the `tips` dataset.

Plot for two Continuous Variables

qplot - Scatter-Plot

```
qplot(PErisk$gdpw2, PErisk$barb2, geom = "point") + geom_smooth(m
```

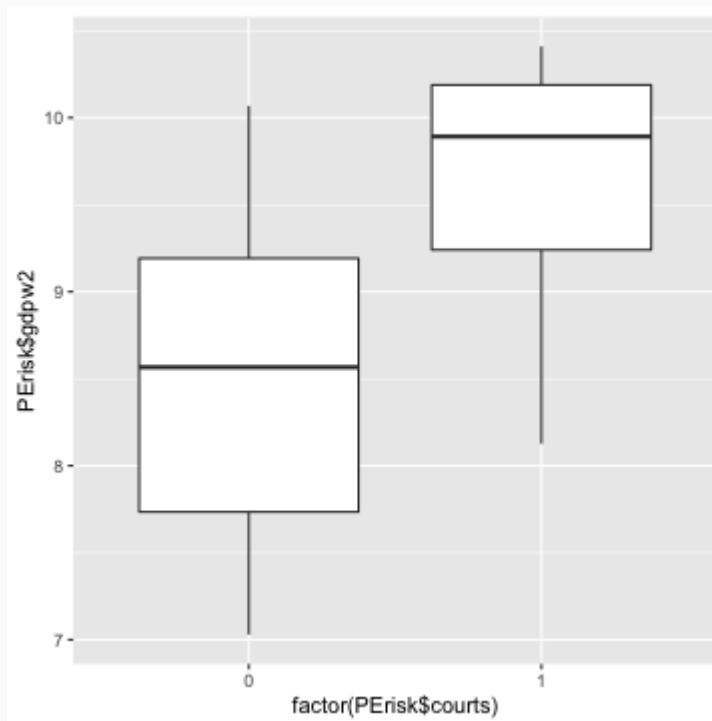
```
## `geom_smooth()` using formula 'y ~ x'
```



Plot for Continuous x Discrete Variables

qplot - Multiple Box-Plots

```
qplot(x = factor(PERisk$courts), y = PERisk$gdpw2,  
      geom = "boxplot")
```



- **Your turn:** make a box-plot of the `tips` variable by `smoker` in the `tips` dataset.

Questions?

See you next week!
