QTM 151

Lab 1 - Intro, GitHub, and ggplot2

Umberto Mignozzetti
Summer

Welcome to QTM 151!

Today's Agenda

About the QTM 151 class

Introductions

GitHub

And `qplot` graphs

QTM151

QTM151

This class will perfect your computational stats skills

We will work with R and R Studio, to teach you:

- To build graphs, maps, interactive webapps
- Data wrangling and high level processing.

All in eight sections!

Let's check the syllabus!

Introductions

Introductions

My name is Umberto Mignozzetti

I am a Political Scientist, with Ph.D. from NYU.

My work concentrates in the field of Comparative Political Economy, where I study how to improve welfare in developing economies.

I am Brazilian, and in my free time I love to watch and play soccer (nowadays mostly in my videogame).

Introductions

What about you?

- What is your name?
- What is your major (current or planned)?
- Tell me something cool about you.

GitHub

GitHub

GitHub is one of the best platforms for saving and retrieving code.

You can save code, check older versions of your own code, and also, build a portfolio that later the firms that will hire you can check.

For this class, I strongly encourage you to use GitHub extensively.

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

Let's check my GitHub!

qplot

qplot

- `qplot`, as the names 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  3.1.2     ✓ dplyr   1.0.7
## ✓ tidyr   1.1.3     ✓ stringr 1.4.0
## ✓ readr   1.4.0     ✓ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse_c
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Loading datasets

```
# Loading tips dataset
tips ← read.csv('https://raw.githubusercontent.com/umbertomig/qtm
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
```

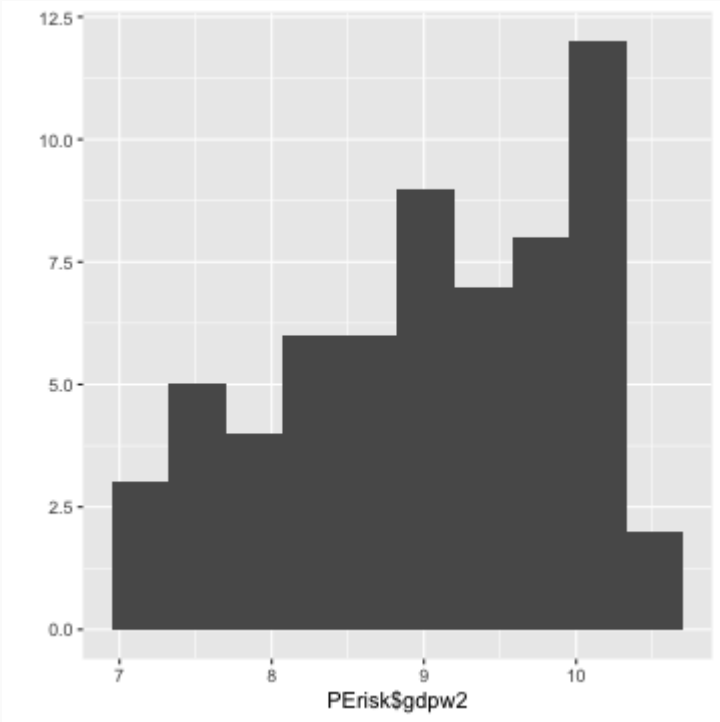
```
# Loading PErisk dataset
PErisk ← read.csv('https://raw.githubusercontent.com/umbertomig/c
head(PErisk, 2)
```

```
##     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

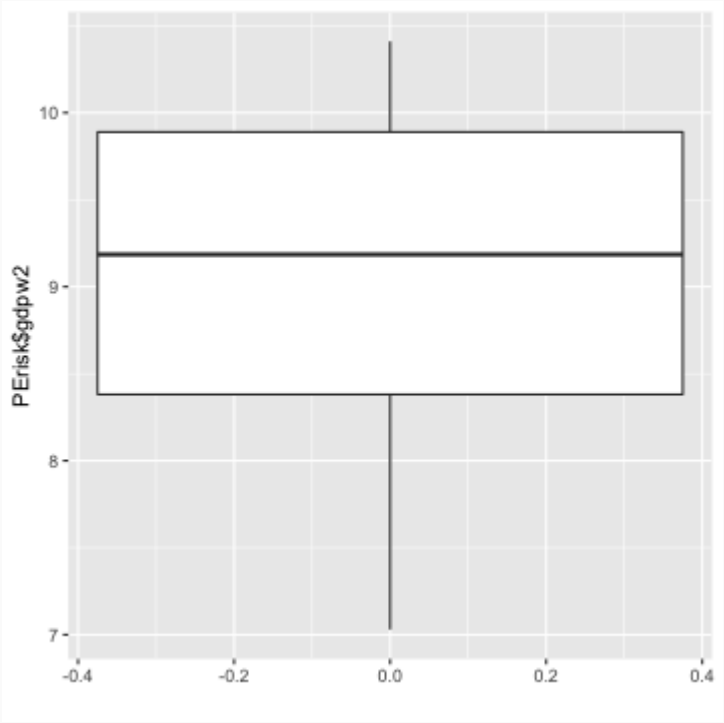
```
qplot(PErisk$gdpw2, geom = "histogram", bins = 10)
```



- **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$gdpw2, geom = "boxplot")
```

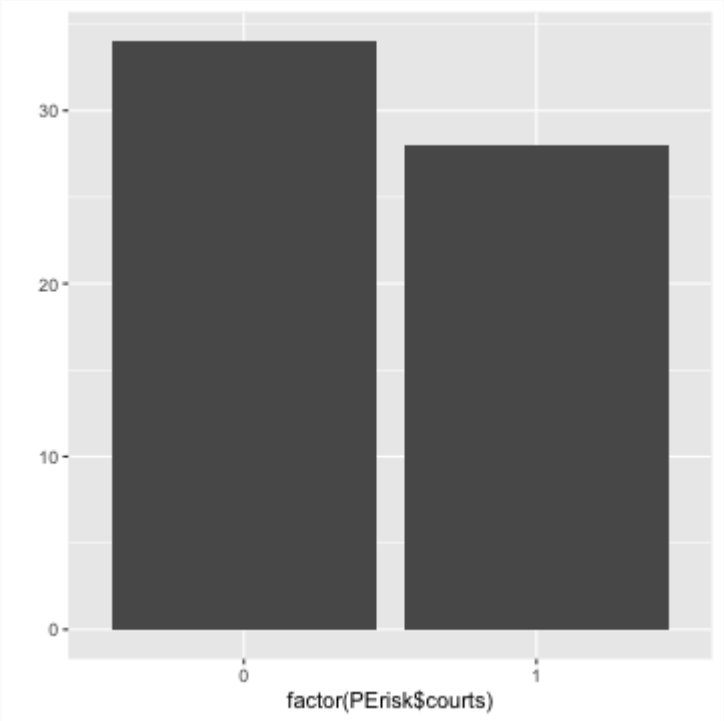


- **Your turn**: make a boxplot of the `totbill` variable in the `tips` dataset.

Plot for Discrete Variables

qplot - Bar-Plots

```
# Bar-plot of courts
qplot(factor(PErisk$courts), geom = "bar")
```



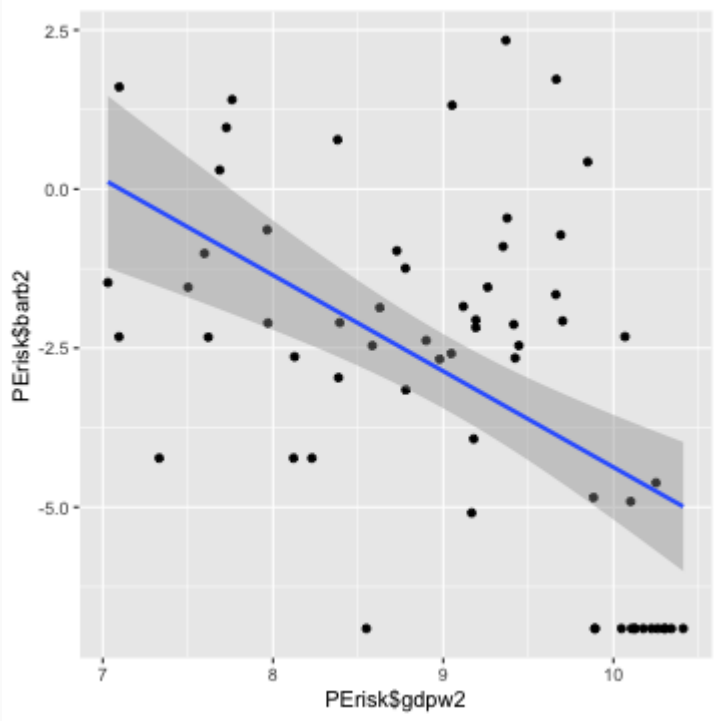
- **Your turn**: make a bar-plot of the `smoker` variable in the `tips` dataset.

Plot for two Continuous Variables

qplot - Scatter-Plot

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

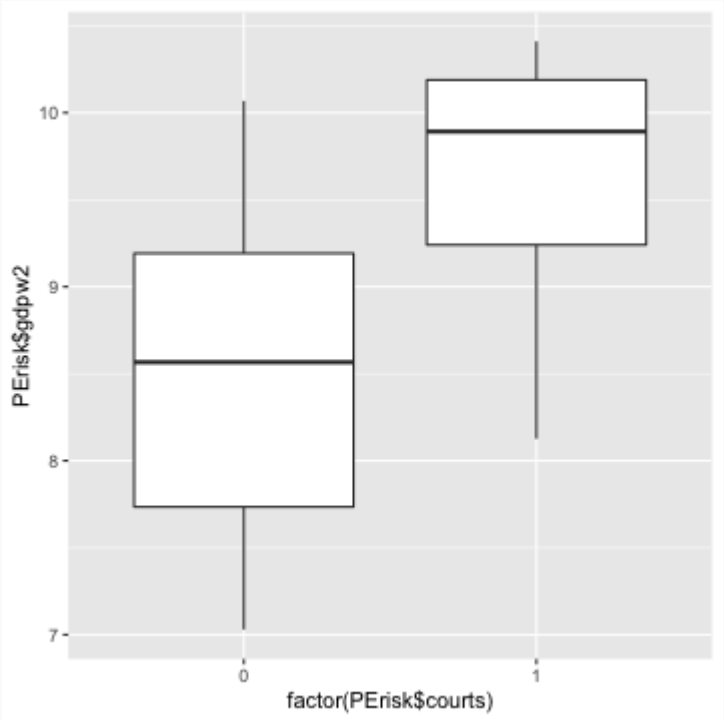
```
## `geom_smooth()` using formula 'y ~ x'
```



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 class!