Data @ Night - What's in a name?

Chris Hua

26 October 2016

About us

- Chris Hua
- Jack Soslow
- Yoonduk Kim

Goal: spend time getting familiar with R and basic data manipulation

Babies!

Figure 1: Hillary

Install packages and load data

Install our easy to use `babynames` package:

```
if(!require(devtools)) {
  install.packages(devtools)
}
devtools::install_github(
    "wharton-data-analytics/babynames")
```

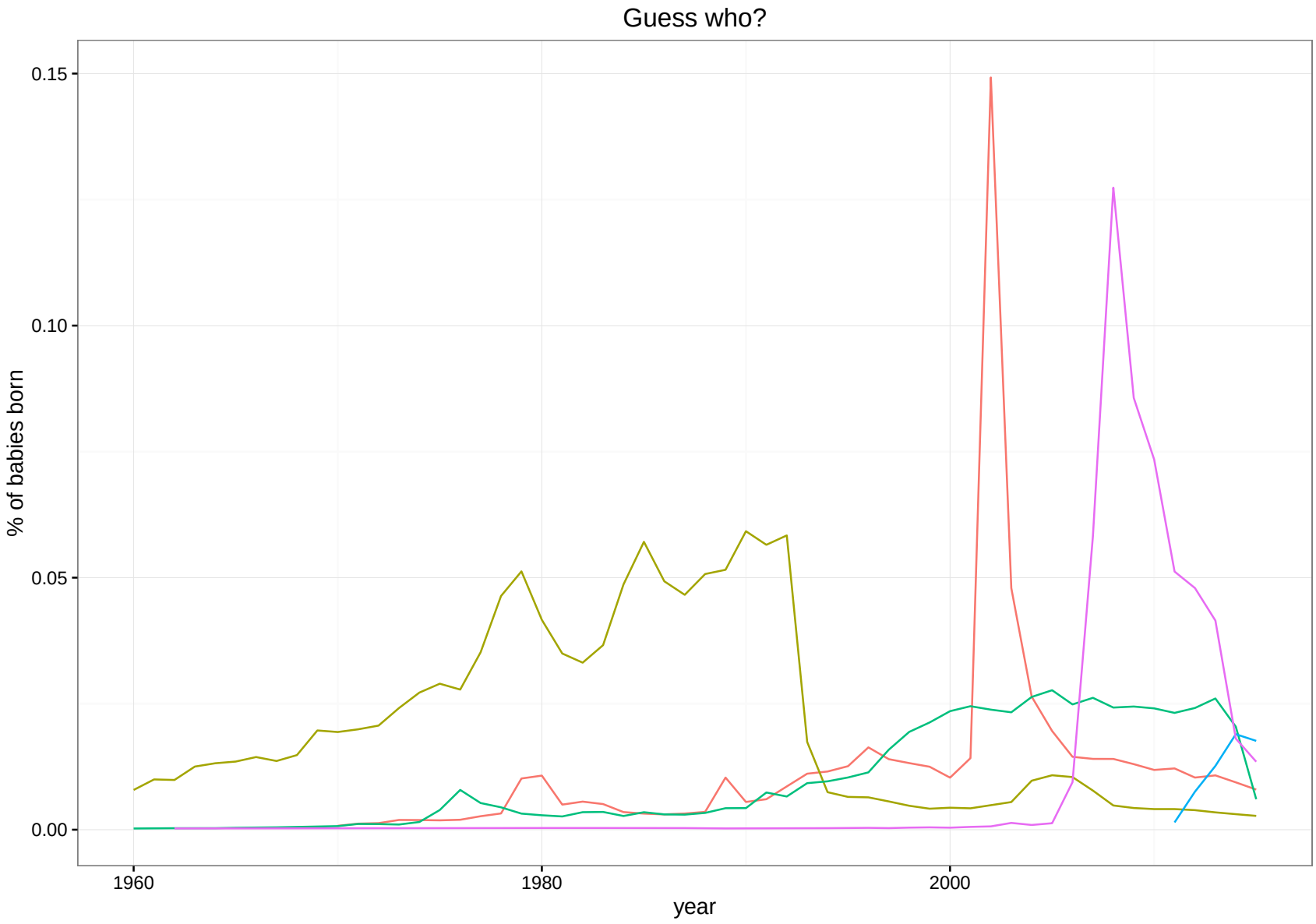
or look in the R code for a work around. Also, install a few more packages...

Explore the data!

Useful commands:

```
head(babynames)
babynames %>% filter(name == "")
```

Trends



Poisoned names

Names that drop off - a lot - between years

How to detect?

"Relative Risk"

Output

year	name	loss
1983	Marquita	-19.14
1957	Tammy	-15.48
1989	Kiara	-10.41
2002	Ashanti	-9.51
1931	Marlene	-7.94
1976	Jaime	-7.51
1995	Alondra	-7.20
1955	Sabrina	-7.00
1979	Brianne	-6.85
1986	Ashton	-6.40
1982	Kayla	-6.05
1991	Iesha	-5.72
1957	Tami	-5.54
1988	Shanice	-5.53
1998	Mya	-5.45

Graph

```
top_names <- names_loss %>%
    top_n(15, wt = loss) %>%
    select(name) %>%
    unique %>% unlist
```

Outliers!