Modern KPI Tracking in R

Nelson Auner

Affirm, Inc

twitter.com/nelsonauner

Why should I?

1. Put important business metrics in the hands of the people who need them

2. Generate communication between teams

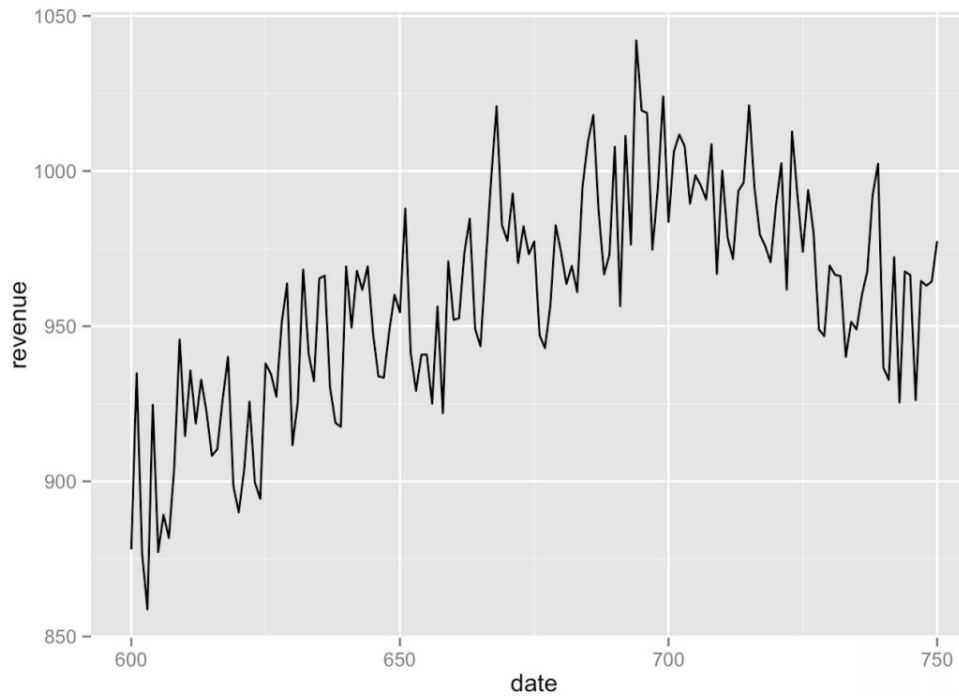3. Reveal patterns or trends you wouldn't have thought to look for

Example

Why not dashboards?

Execs don't want to click to links on dashboards

They don't want to open an attachment

They want to read it on their iphone!

It really works!

"I absolutely *love* the attention that metrics are getting here"

- Tech CEO, in the midst of a company-wide email discussion in response to weekly metrics email

The Recipe

1. Connect R with your data warehouse

1. Formulate analysis

1. Set up ELT job, actual email delivery

1. Connect R with your data warehouse

```
flowers <- src_postgres()
```

```
connect_to_my_db  <- function(searchpath="public") {
    connection <- src_postgres(host="localhost",port="5439",dbname="my_db",
                               user = getOption("ssh_cred")['user'],
                               password =getOption("ssh_cred")['password'],
                               options = paste0("-c search_path=",searchpath))
    return(connection)
}
```

Loading in data

- dplyr syntax

```
flowers %>%
  tbl("iris") %>%
  summarise(meanSepal = mean(Sepal.Length),
            meanPetal = mean(Petal.Length))
```

- You can also use SQL syntax

```
flowers %>% tbl(sql("SELECT * FROM iris"))
```

2. Formulate Analysis

This week vs. last week
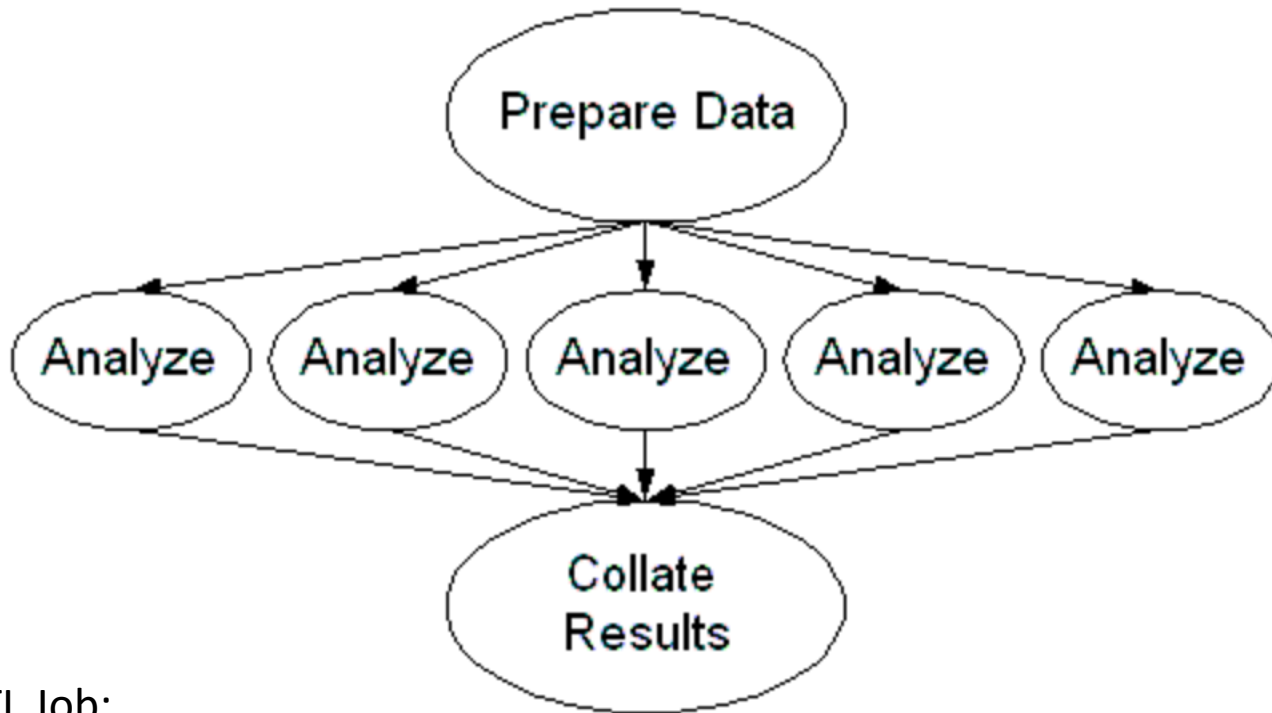
This week vs. projected goals

Flower statistics

```
flowers %>%
  tbl("iris") %>%
  summarise(meanSepal = mean(Sepal.Length),
            meanPetal = mean(Petal.Length)) %>%
  collect()
```

```
Source: local data frame [1 x 2]

  meanSepal meanPetal
      (dbl)     (dbl)
1  5.843333     3.758
```

3. DAG Job + Email delivery

For ETL Job:
CRON job
Luigi
Airflow

For automated email:
mailR

Q&A

Connect:

nelson@nelsonauner.com

twitter.com/nelsonauner