**Data Source**

[Data about TED Talks on the TED.com website until September 21st, 2017](https://www.kaggle.com/rounakbanik/ted-talks/data)

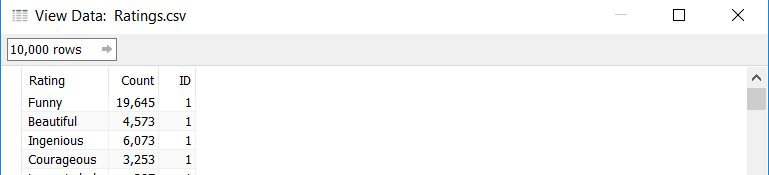
**Data preparation**

* Use python to split ratings column, and create separate rating excel file

Original rating data looks like this:



We were able to split it into this:



Codes:

df['ratings'][0]

rating\_names = set([])

*# get a list of rating "names".*

def split\_ratings(ratings):

val = ast.literal\_eval(ratings)

for rating **in** val:

rating\_names.add(rating['name'])

df['ratings'].apply(split\_ratings)

def get\_count\_from\_ratings(ratings, col\_name):

val = ast.literal\_eval(ratings)

for rating **in** val:

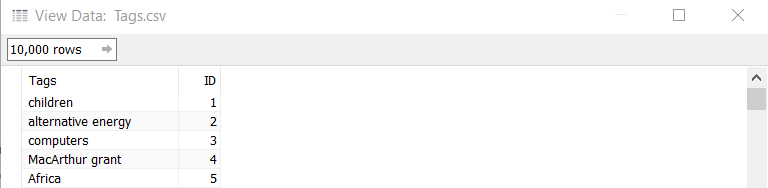
if rating['name'] == col\_name:

return rating['count']

for name **in** rating\_names:

df[name] = df['ratings'].apply(lambda rating : get\_count\_from\_ratings(rating, col\_name=name))

* Change film\_date column to correct format
  + <Film date> originally in unix stamp format
  + Create a calculated field, save as Date:
    - DATEADD('second', INT([Film Date]), #1970-01-01#)
* Extract major speaker occupation (for each speaker, limit the occupation to 1)
  + Through splitting by the first “;” in <Speaker Occupation>
* Split tags column through splitting by the first ”,” and create tags excel file

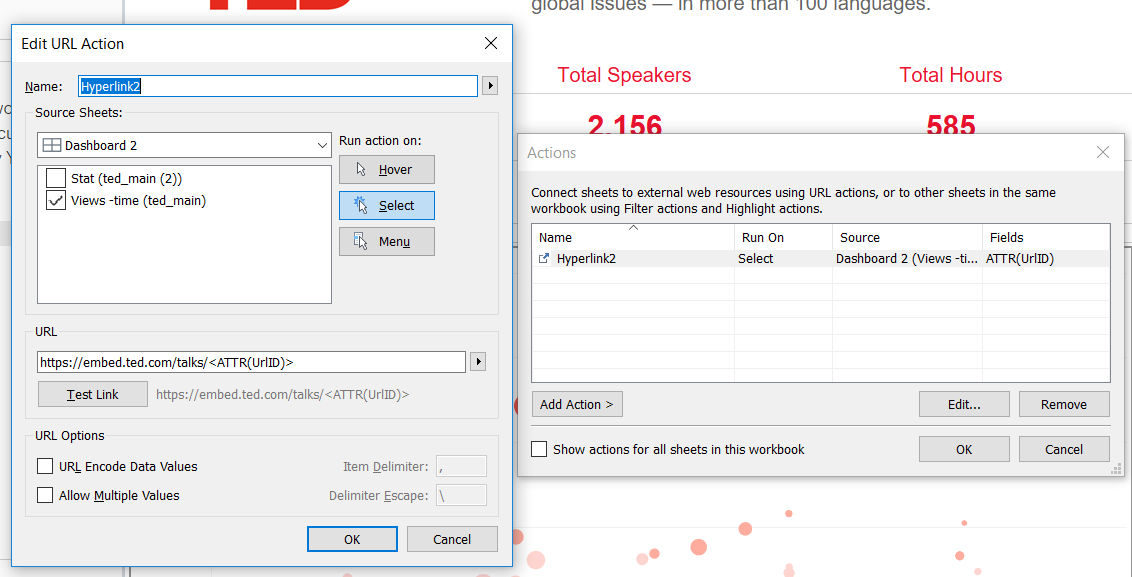


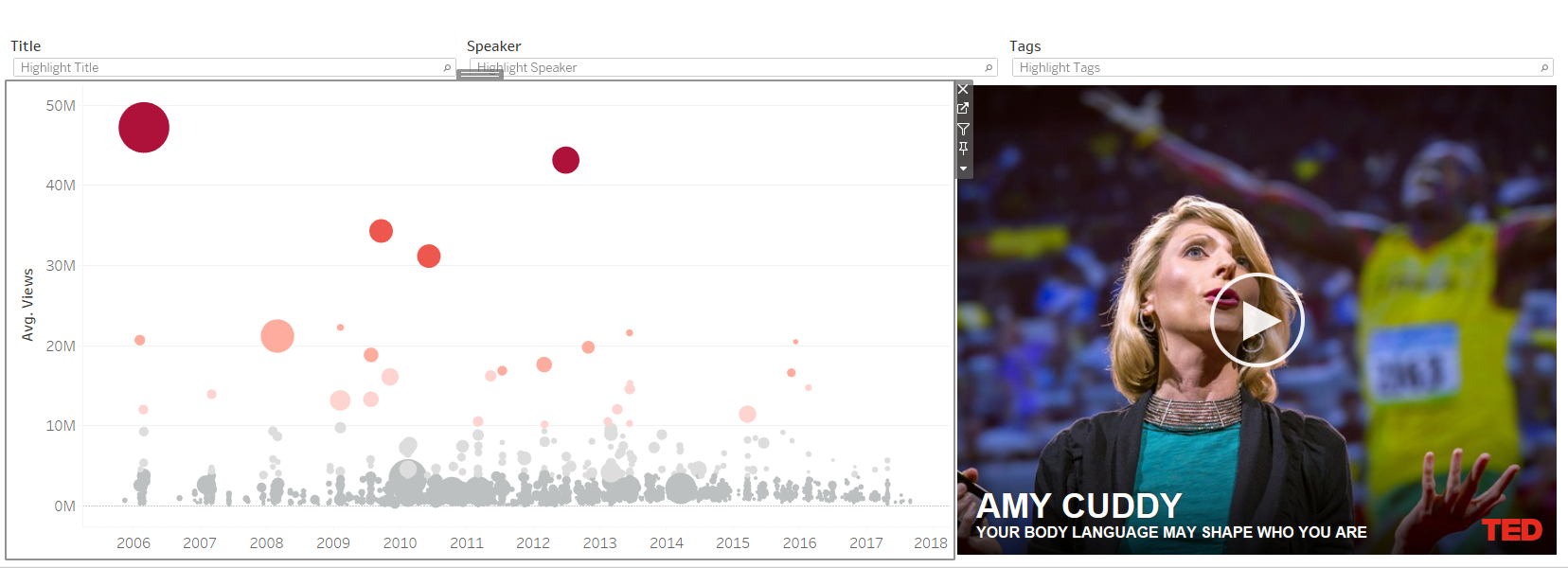
* Union ted\_main, rating and tag tables to form data source
* Split <Url> by the last “/” to get <UrlID>

**Worksheets**

Views vs time worksheet

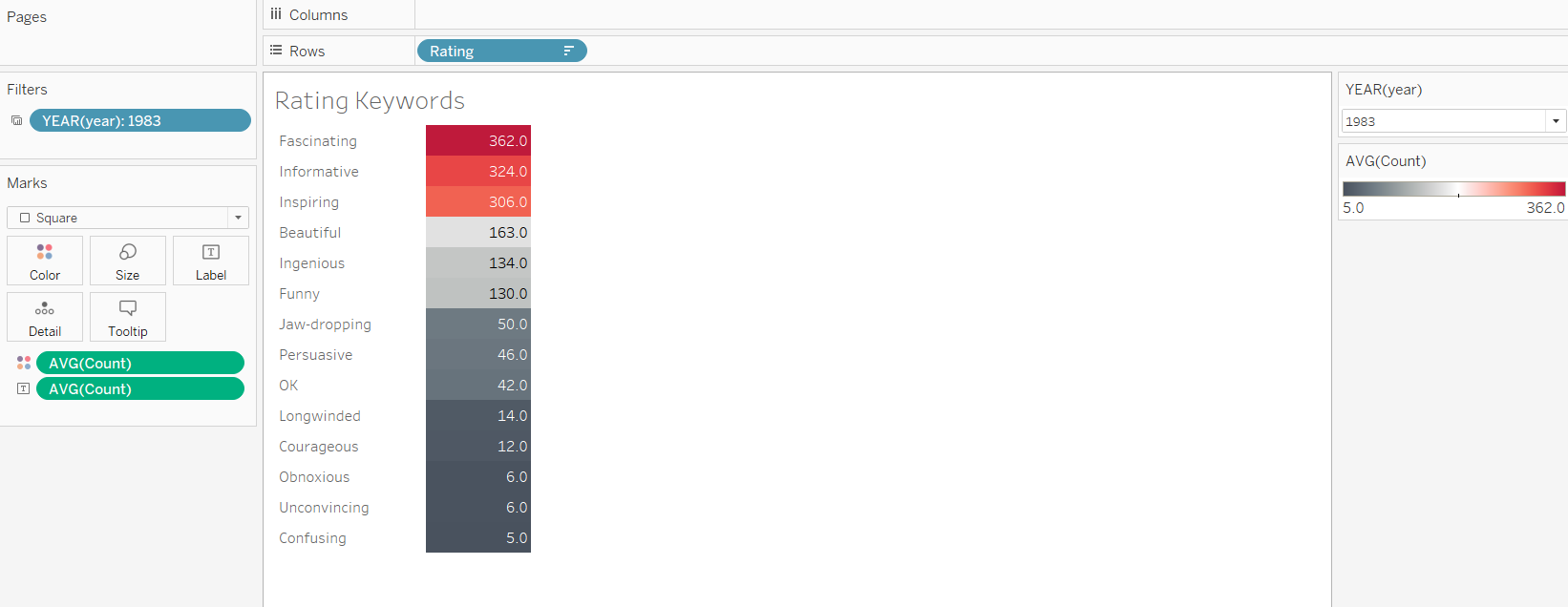
1. Date into columns, change to day, continuous
2. Views into rows, change to average
3. Views into marks, change to average, color
4. Change color to Red-Black-White Diverging, opacity 78%, 8 steps color, reversed and centered
5. Name into marks, change to detail
6. Title into marks, change to detail, show highlighter
7. Tags into marks, change to detail, show highlighter
8. Speaker into marks, change to detail, show highlighter
9. Embed video into dashboard:
   * + - 1. Drag <UrlID> to “Tooltip” so that it becomes an attribute that we can use in embedding videos
         2. Go to a dashboard, Insert a “web page” space
         3. “Dashboard” 🡪 Actions 🡪 leave the window blank and click OK 🡪 Add Action
         4. Choose the sheet we want to add videos on 🡪 choose “On Click” on the right
         5. Edit the link
         6. Enter [https://embed.ted.com/talks/ATTR<UrlID>](https://embed.ted.com/talks/ATTR%3cUrlID%3e)





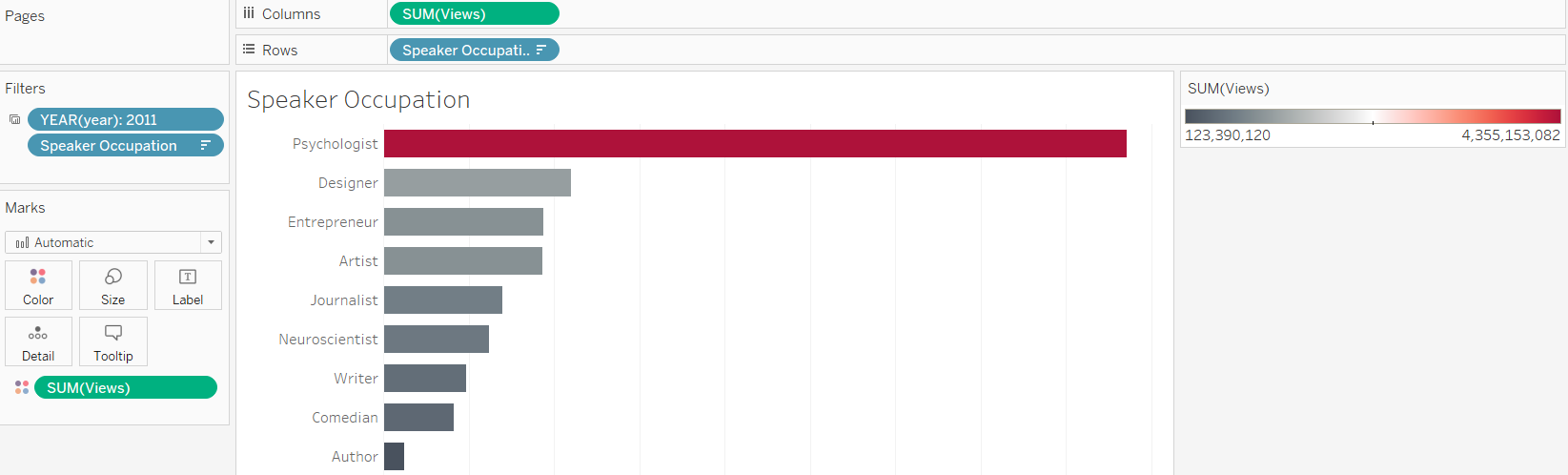
Rating keywords worksheet

1. Rating into rows
2. Average count into marks, add color and label
3. Change color to Red-Black-White Diverging, reversed, centered at 190
4. Sort order descending by average count
5. Show title
6. Tooltip shows rating and avg count



Speaker Occupation worksheet

1. Views into columns, change to SUM
2. Speaker Occupation into rows
3. Views into marks, change to SUM, color
4. Change color to Red-Black-White Diverging, reversed
5. Speaker Occupation into filter, show only top 10 by views SUM
6. Sort order descending by sum views
7. Show title
8. Tooltip shows speaker occupation and sum views



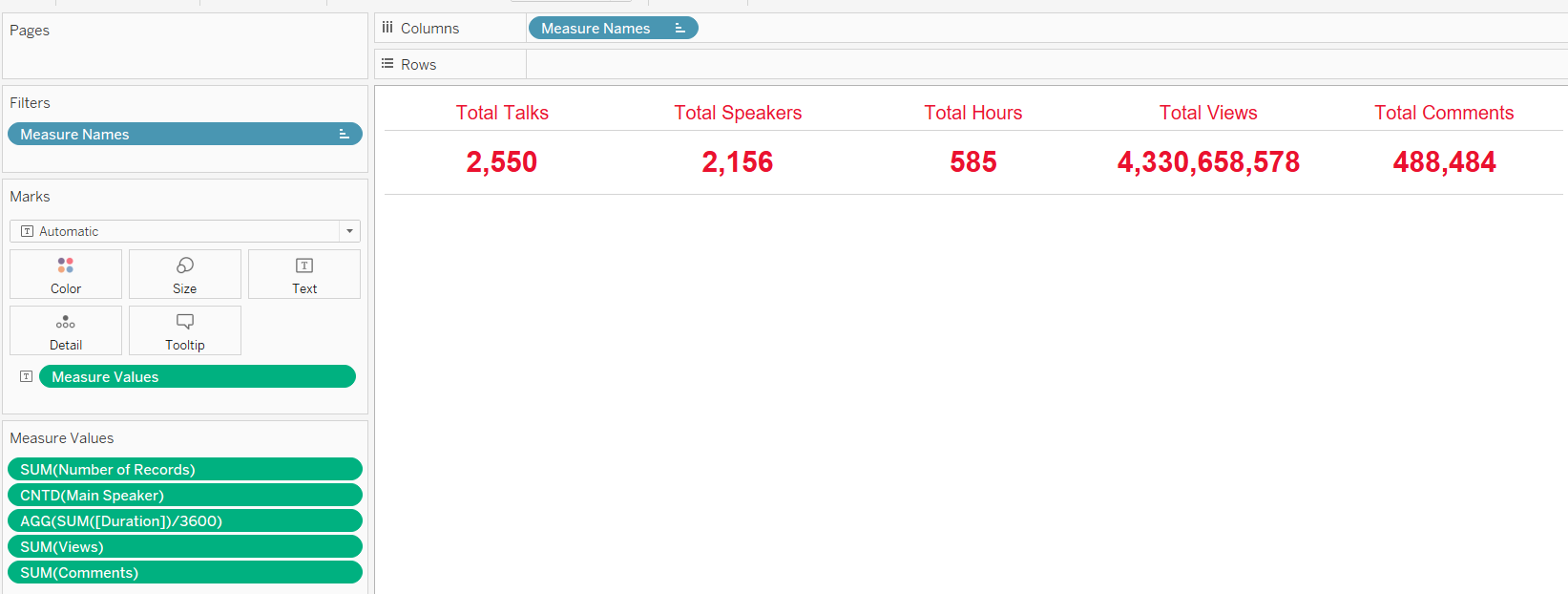
Top Tags by Year worksheet

1. Tags into Marks, change to color
2. Number of Records into Marks, change to size, SUM
3. Tags into Marks, change to label
4. Color change to Gray Warm
5. Tags into filter, show top 25 by sum number of records
6. Year into pages
7. Tooltip shows tags, year and number of records
8. Apply year filter for all three worksheets above.



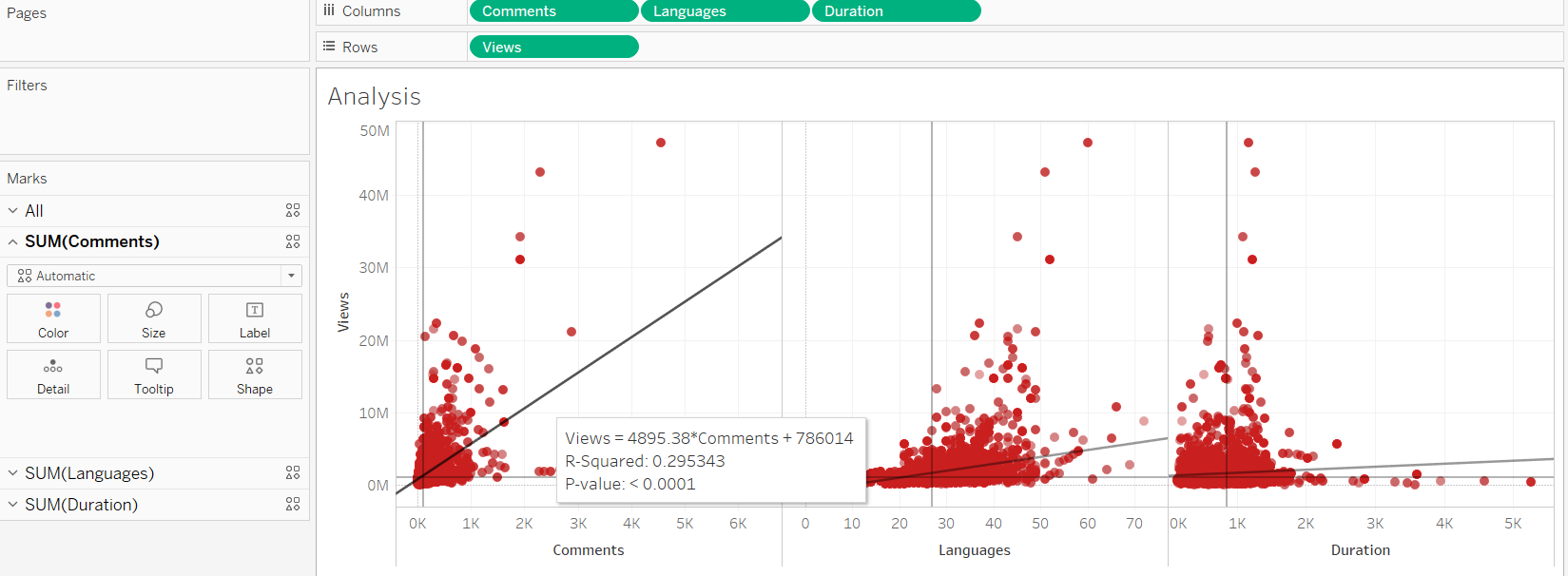
Stat worksheet

1. Measure Values into marks, change to label and select:
2. Sum number of records
3. Count distinct main speaker
4. Average of sum of duration, divided by 3600 (convert from seconds to hours)
5. Sum comments
6. Sum views
7. Measure Names into filters, choose total comments, total hours, total speakers, total talks and total views
8. Measure Names into columns
9. Change font to Arial 12pt, red
10. Change view to fit width



Analysis worksheet

1. Comments into columns
2. Languages into columns
3. Duration into columns
4. Change all three column elements to shapes
5. Views into rows
6. Selected color HTML code: #aa0000
7. Add trend line and median with 95%CI from analytics tab



Selected ted talk worksheet

* Title, Speaker, Event, Date, Speaker Occupation into rows
* Title, Speaker, Event, Date, Speaker Occupation into filters, sorted by avg views
* Avg views, comments, duration and languages into marks, change to color
* Description into marks, change to ATTR in tooltip
* Go to dashboard, choose action, add hyperlink and put <ATTR(Url)> in the URL

