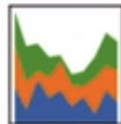
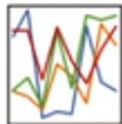


NumPy

pandas

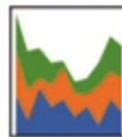
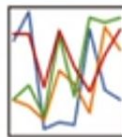
$$y_{it} = \beta^t x_{it} + \mu_i + \epsilon_{it}$$



matplotlib

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



Data Frames



Data Visualization

SIMPLE PLOTS

SMALL AMOUNT OF CODE

PROFESSIONAL LOOK



Kids

DVD



Search



Tom

Top Picks for Tom



The OA

93% Match 2016 TV-MA 1 Season

To some, she's a miracle. To others, a freak. Only a few understand where she's been: to the brink of an unseen world.

Like, Dislike, Add buttons



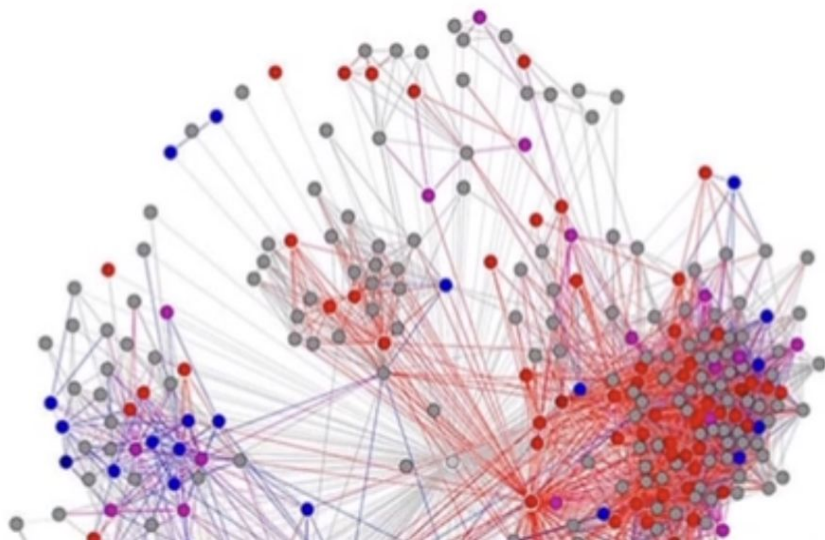
Irreverent TV Shows



MAY 7, 2015

Exposure to Diverse Information on Facebook

By: Eytan Bakshy, Lada Adamic, Solomon Messing





Sponsored products you might like

						
\$295.80	\$149.00 list price \$229.00 save \$80.00	\$49.89	\$6.98	\$329.00 list price \$399.00 save \$70.00	\$279.00	\$249.00
Acer Sparkly Silver 14" CB3-431-C5FM...	14" Full HD IPS screen Laptop, Ultra-Slim...	Keurig® Classic Series Chili Red Keurig® K15...	Starbucks Espresso Roast Dark Ground...	Acer Aspire ES 15 Series ES1-572-33BP 15.6"...	HP 14-ak040wm 14" Chromebook,...	Dell Inspiron 11 3168 11.6" Laptop,...
★★★★★ 5	★★★★★ 86	★★★★★ 262	★★★★★ 53	★★★★★ 2	★★★★★ 2	★★★★★ 3116
2-Day Shipping	2-Day Shipping	2-Day Shipping	2-Day Shipping	2-Day Shipping	2-Day Shipping	2-Day Shipping



Top-selling Pickup Discount items





learning has shown promise in multiple fields but requires considerable computing power.

Fortunately for Google, this is one area in which it is well equipped.

The search giant applied its large-scale neural network training system to the work. Google built the system to train networks of tens of thousands of CPU cores to perform a task. In the drug discovery project, the training entailed equipping the network to comb through 37.8 million data points covering more than 200 different biological processes. After running the system for more than 50 million CPU hours, Google has concluded the inclusion of data from multiple sources allows it to make more accurate predictions of the efficacy of a drug across different diseases.

Even greater scales are in Google's sights. At the time of writing a paper on the project, Google had scaled the system up to 239 tasks and the upward efficiency trend was yet to plateau. Similarly, the addition of more data was found to increase efficiency, too. The researchers have cast lustful looks at the "vast private stores of experimental measurements" locked away at Big Pharma companies as they try to figure out the next steps for the model. More data and more tasks are the near-term goals for the project.

Whether the efficiency gains touted by Google will have an effect on drug discovery remains to be seen. Even the paper's authors accept that the complexity of drug discovery could limit the impact of the approach, but overall are as optimistic as one would expect Google staffers to be about the potential for data and algorithms deliver improvements.

- read the [paper](#) (PDF)
- check out Google's [blog post](#)
- and [VentureBeat's take](#) on controversy

FREE BIOTECH NEWSLETTER

Join over 150,000 subscribers who benefit from FierceBiotech's coverage on such topics as biopharma news and deals, clinical trials, FDA decisions and more.

Join for free

Email

Sign up



September 25-27, 2017

Renaissance Boston Waterfront Hotel
Boston, MA

Born	George William James			
	October 5, 1949 (age 67)			
	Wilton, Kansas, U.S.			
3 Days	4 Tracks	300+ Attendees	4 In-Depth Workshops	10+ Hours of Networking
Known for				



IP[y]:
IPython



WRANGLING

▼ Gather

▼ Assess

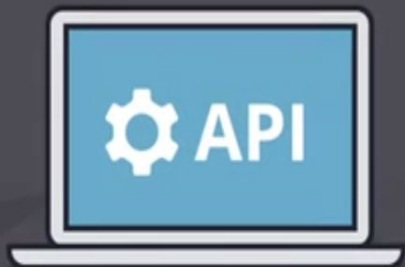
▼ Clean

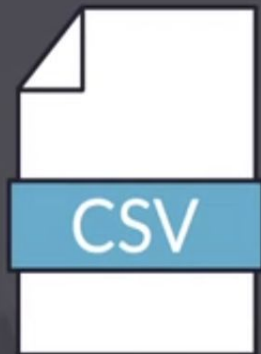
EDA

▼ Explore

▼ Augment

GATHERING





COMMA SEPARATED VALUES

cancer_data.csv - Google Sheets

Secure | <https://docs.google.com/spreadsheets/d/1IE7WJbMogcaJ8RhkenC3j6T>

cancer_data.csv

File Edit View Insert Format Data Tools Add-ons Help Last edit was 31 minutes ago

fx id

	A	B	C	D	E	F
1	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean
2	842302	M	17.99	10.38	122.8	100
3	842517	M	20.57	17.77	132.9	132
4	84300903	M	19.69	21.25	130	120
5	84348301	M	11.42	20.38	77.58	386
6	84358402	M	20.29	14.34	135.1	129
7	843786	M	12.45	15.7	82.57	477
8	844359	M	18.25	19.98	119.6	104
9	84458202	M	13.71	20.83	90.2	577
10	844981	M	13	21.82	87.5	519
11	84501001	M	12.46	24.04	83.97	475
12	845636	M	16.02	23.24	102.7	797
13	84610002	M	15.78	17.89	103.6	78
14	846226	M	19.17	24.8	132.4	112
15	846381	M	15.85	23.95	103.7	782
16	84667401	M	13.73	22.61	93.6	578
17	84799002	M	14.54	27.54	96.73	658
18	848406	M	14.68	20.13	94.74	684
19	84862001	M	16.13	20.68	108.1	798
20	849014	M	19.81	22.15	130	126

cancer_data.csv

cancer_data.csv -- ~/Downloads

cancer_data.csv

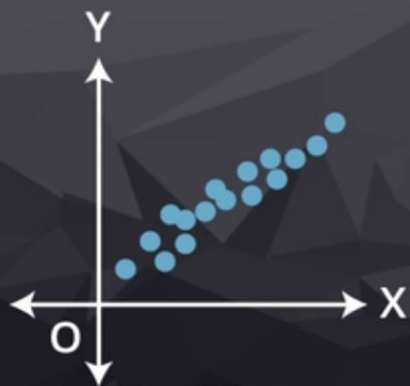
```

1 id,diagnosis,radius_mean,texture_mean,perimeter_mean
2 842302,M,17.99,10.38,122.8,1001.0,0.1184,0.2776,0.3
3 842517,M,20.57,17.77,132.9,1326.0,0.08474,0.07864,0.
4 84300903,M,19.69,21.25,130.0,1203.0,0.1096,0.1599,0.
5 84348301,M,11.42,20.38,77.58,386.1,0.1425,0.2839,0.
6 84358402,M,20.29,14.34,135.1,1297.0,0.1003,0.1328,0.
7 843786,M,12.45,15.7,82.57,477.1,0.1278,0.17,0.1578,
8 844359,M,18.25,19.98,119.6,1040.0,0.09463,0.109,0.1
9 84458202,M,13.71,20.83,90.2,577.9,0.1189,0.1645,0.0
10 844981,M,13.0,21.82,87.5,519.8,0.1273,0.1932,0.185
11 84501001,M,12.46,24.04,83.97,475.9,0.1186,0.2396,0.
12 845636,M,16.02,23.24,102.7,797.8,0.08206,0.06669,0.
13 84610002,M,15.78,17.89,103.6,781.0,0.0971,0.1292,0.
14 846226,M,19.17,24.8,132.4,1123.0,0.0974,0.2458,0.2
15 846381,M,15.85,23.95,103.7,782.7,0.08401,0.1002,0.0
16 84667401,M,13.73,22.61,93.6,578.3,0.1131,0.2293,0.2
17 84799002,M,14.54,27.54,96.73,658.8,0.1139,0.1595,0.
18 848406,M,14.68,20.13,94.74,684.5,0.09867,0.072,0.07
19 84862001,M,16.13,20.68,108.1,798.8,0.117,0.2022,0.1
20 849014,M,19.81,22.15,130.0,1260.0,0.09831,0.1027,0.
21 0510126,M,12.54,14.26,87.46,556.2,0.00770,0.00120,0.

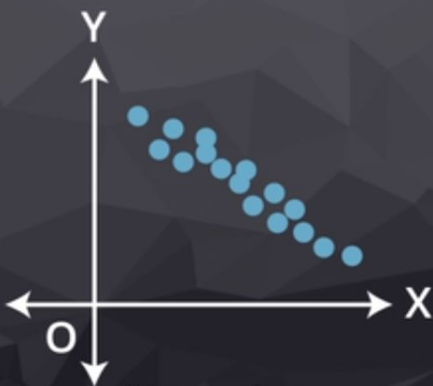
```

cancer_data.csv 1:1

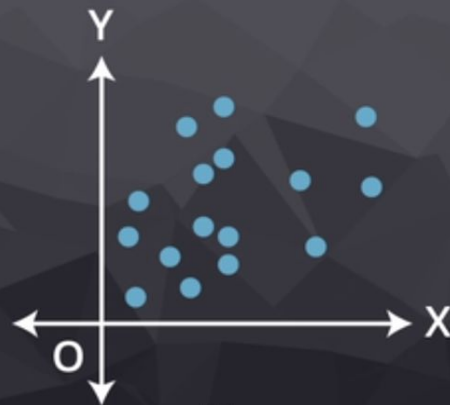
LF UTF-8 Plain Text 0 files



Positive
Correlation

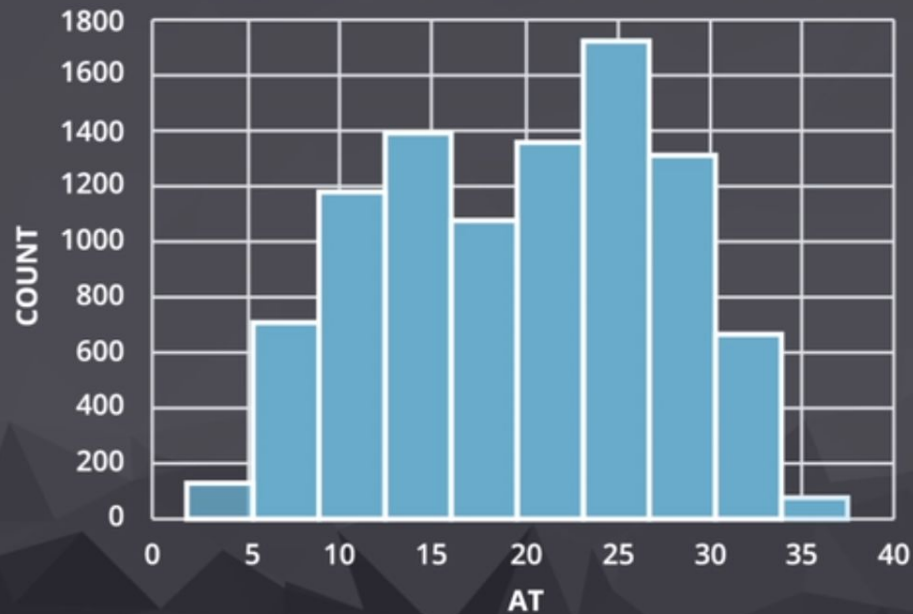


Negative
Correlation

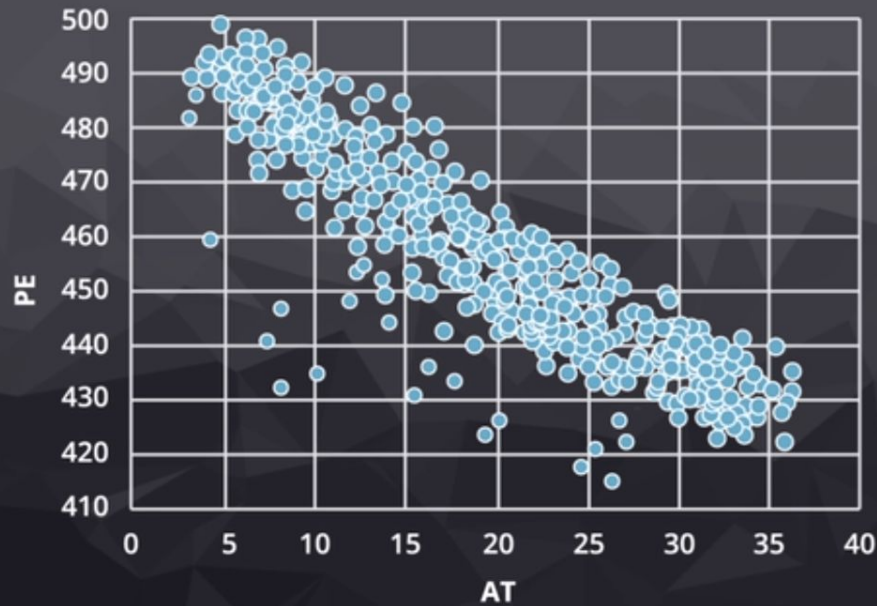


No
Correlation

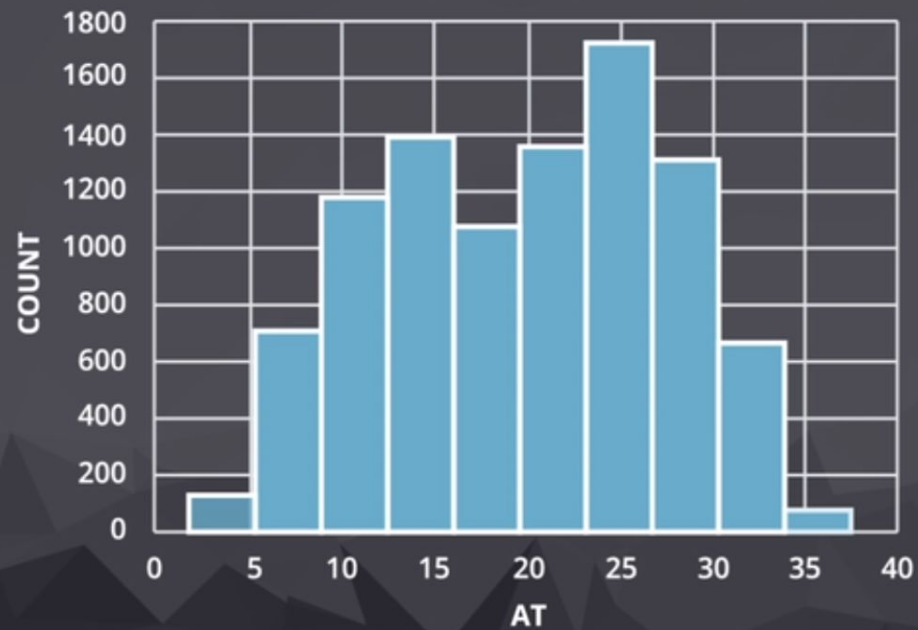
DISTRIBUTION OF TEMPERATURES



PE vs. AT



DISTRIBUTION OF TEMPERATURES



PE vs. AT

