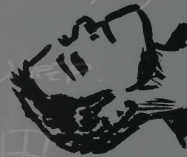


Vince
Croft

SGD Topics #1: Multivariate Visualisation

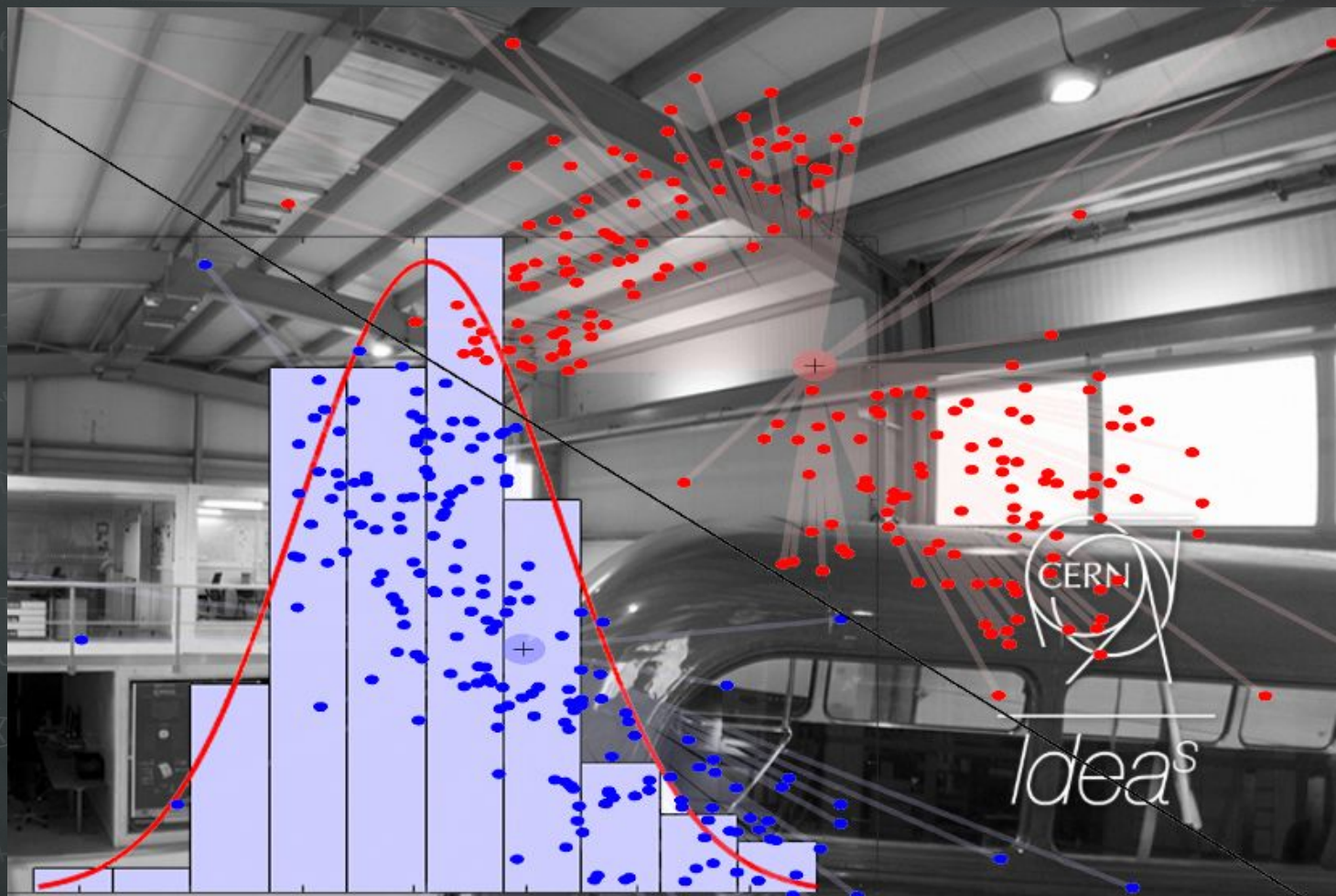
@vincecroft

@DrYacineHaddad



Welcome!

- Hi! thanks to IdeaSquare!



Changes on the Way!

- SGD is on the way
- Two new meetups on the horizon
- Website under production



Société Genevoise des Données

Who are we

Editions

Meetups

Contact

Geneva Society of Data

The home for data analysis, engineering
and machine learning in the Genève region.



Check out our next meetup.

[Sign up now >](#)

Hands On Code

Code describing how to get started understanding modern topics in data is hosted on GitHub. Here we share work, discuss changes, and get feedback in one place to expand our knowledge and experience.

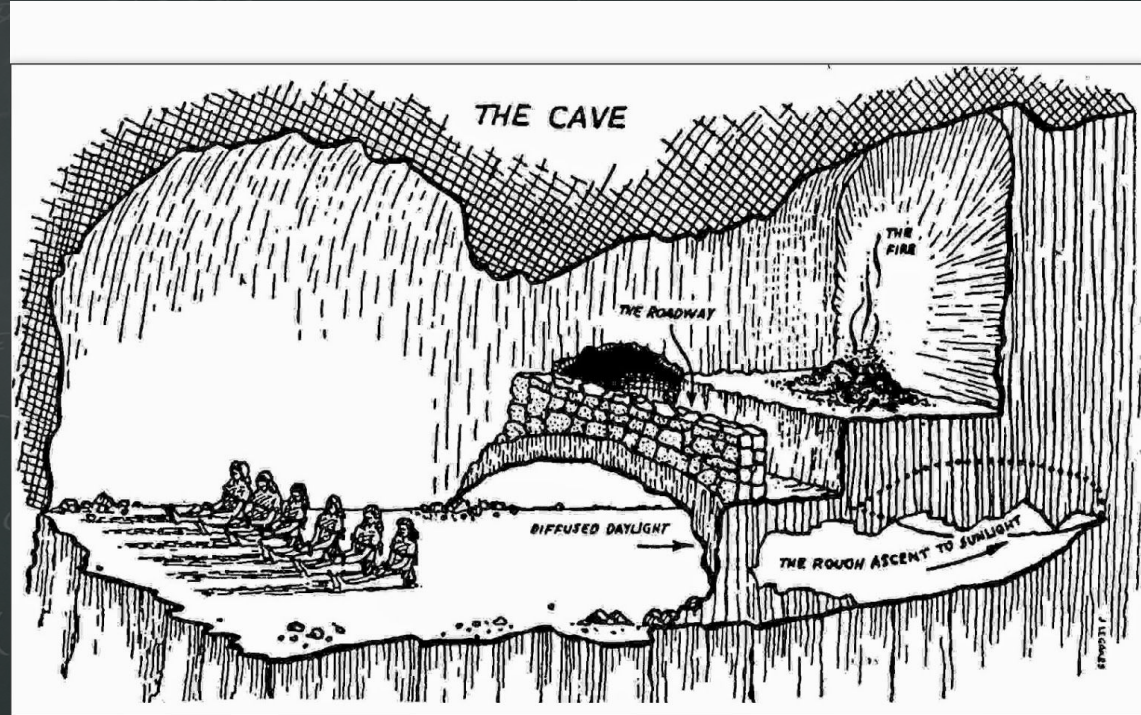


Experiment with new ideas

The society represents a cross-section of data science professionals from all corners of academia and industry a breeding ground for exciting new perspectives on your data.

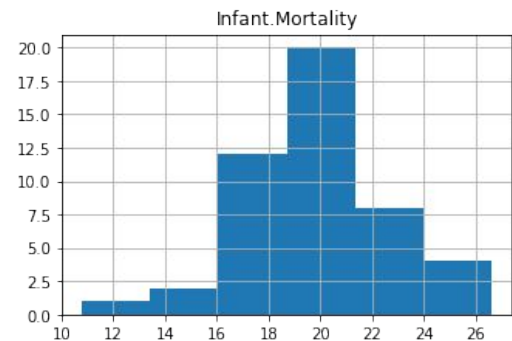
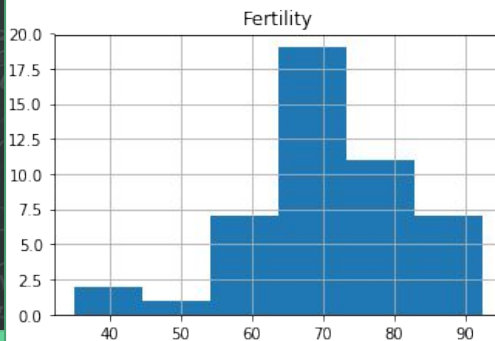
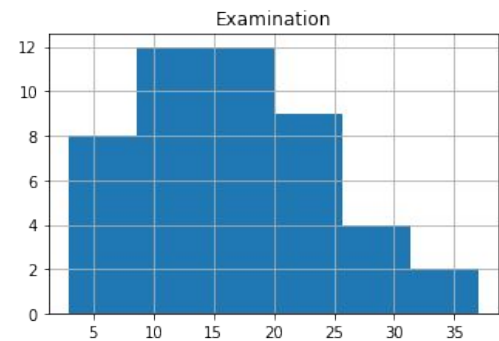
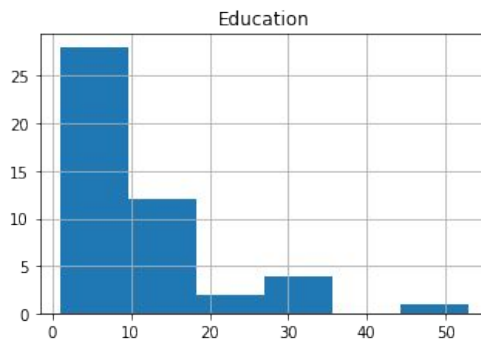
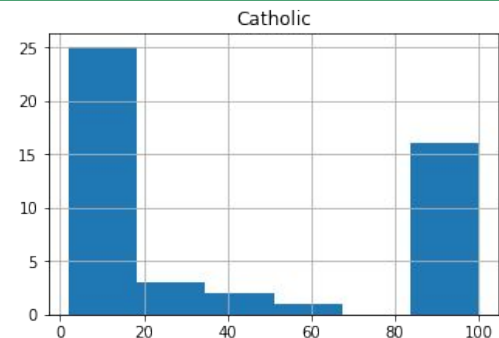
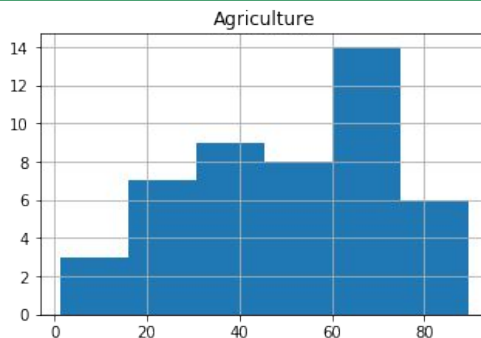
Visualising Data!

- The most important step in dealing with data.
- Continuous, categorical, binned/bucketed, ordered, etc
- What and why?
- Exploring data - EDA



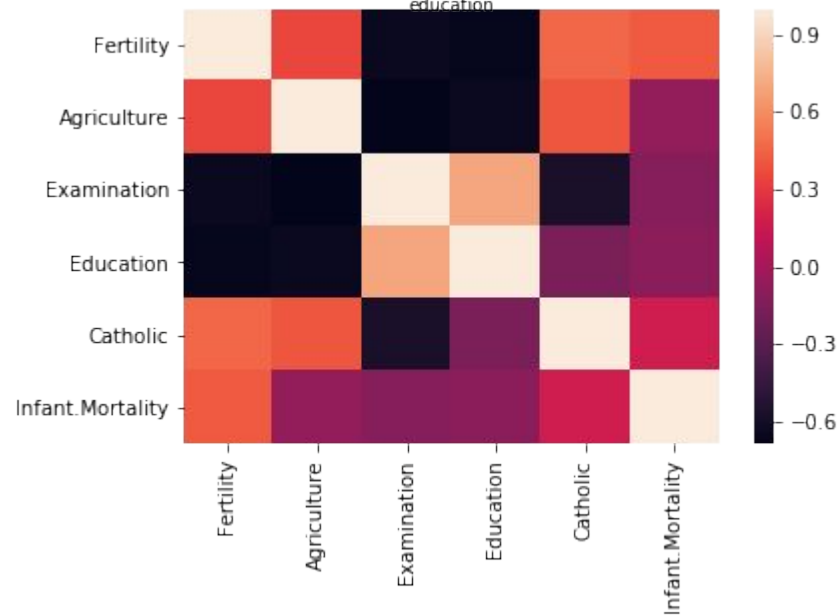
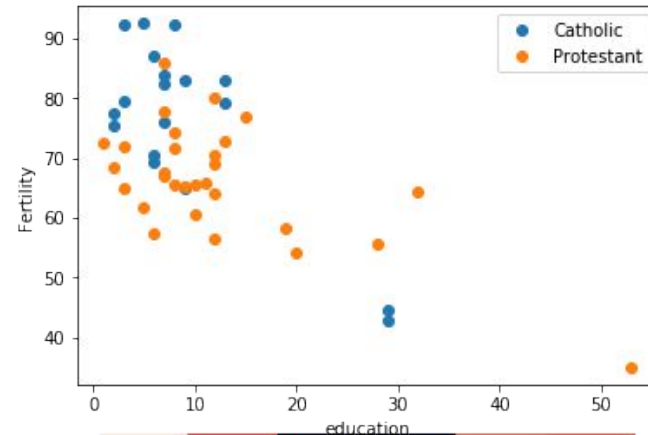
1D - Histograms and Summary data

- Data should be in reference to more data.
- Histograms let us visualise frequency and scale of observations
- Mean, standard deviation, etc let us understand a data set without seeing it all



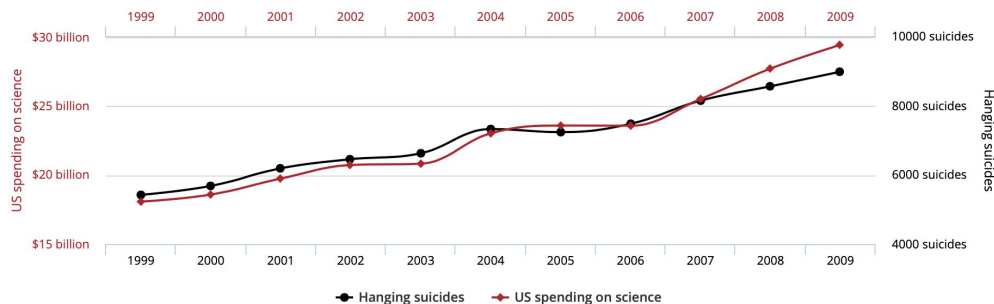
Correlation

- Picturing the relationship between two variables can give the most insight into how you data behaves
- Correlation does not imply causation



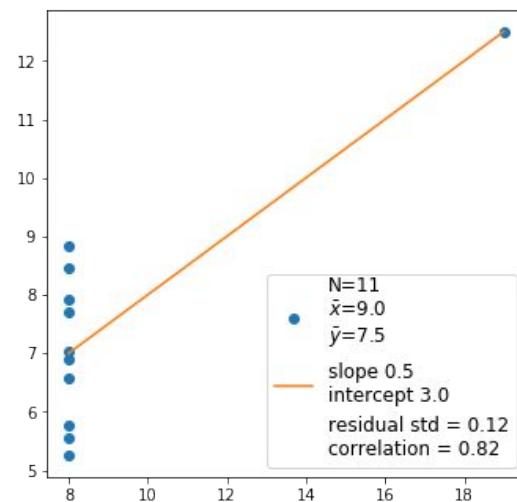
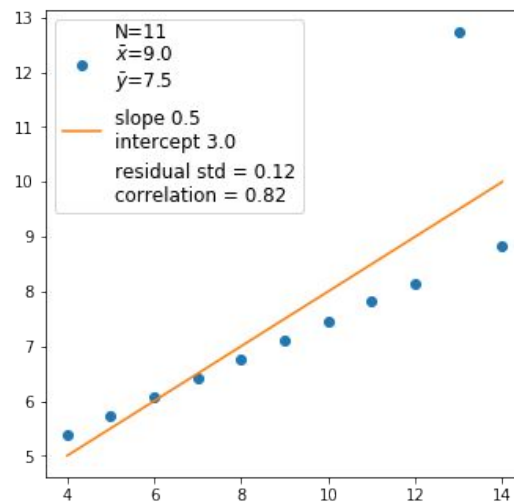
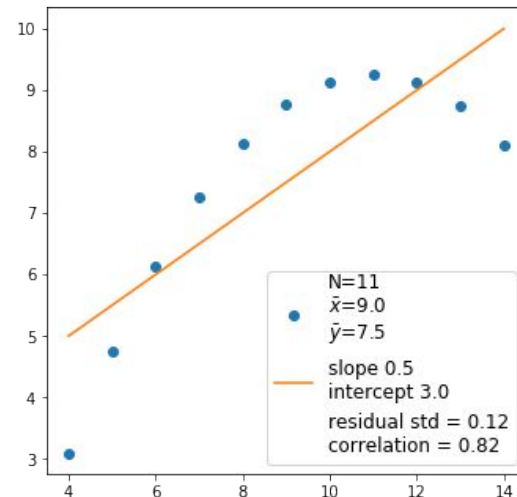
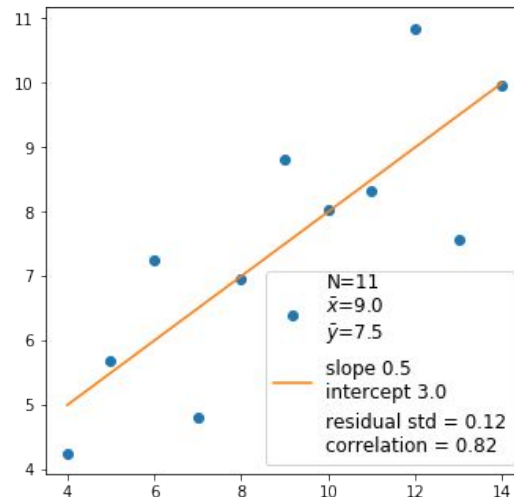
US spending on science, space, and technology correlates with Suicides by hanging, strangulation and suffocation

Correlation: 99.79% ($r=0.99789126$)



A Picture Tells a Thousand Words

- Picturing the data is sometimes the only way to understand it
- The anscombe data sets have identical numbers of points
lines of best fit
standard deviations and correlations



Try it yourself!

<https://github.com/SocieteGenevoiseDonnees/MultivariateVisualisation>

opendata.swiss

Data

Organizations

Applications

About ▾



Find Swiss Open Government data

Learn more about opendata.swiss

6,385
Datasets

Search datasets...



Access the data catalogue using the API