

DATA SCIENCE

11 WEEK PART TIME COURSE

Week 2 – Data Visualisation
Monday 14th December

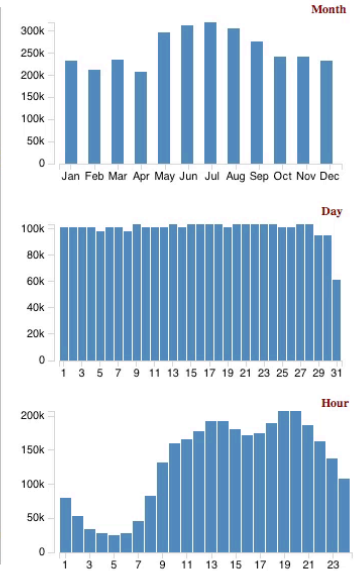
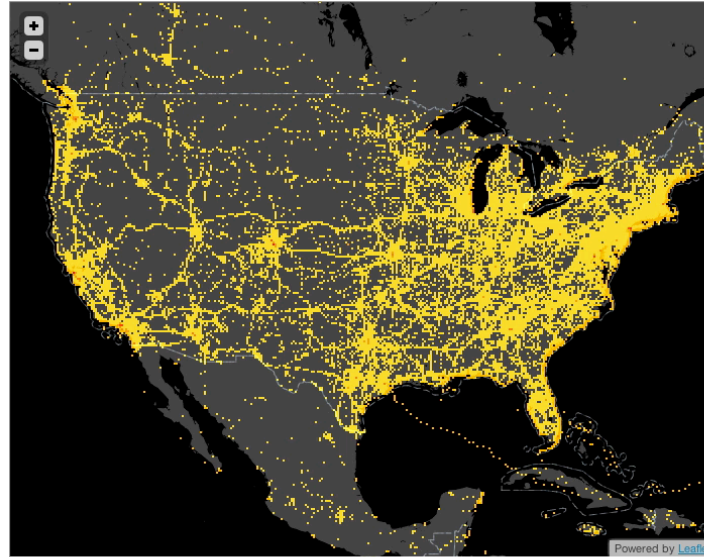
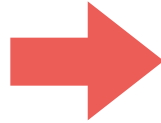
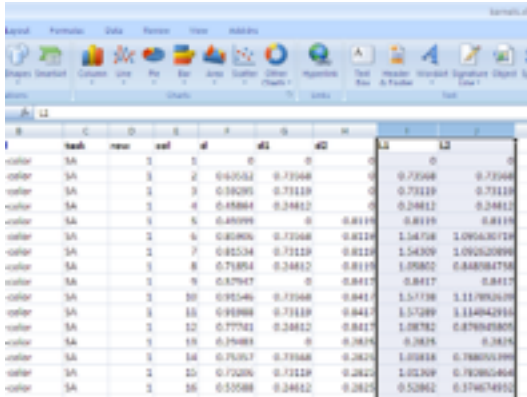
1. Guest Speaker !!?!?!?
2. What is Data Visualisation?
3. Why do we visualise data?
4. How do we visualise data?
5. Lab
6. Discussion

DATA SCIENCE PART TIME COURSE

WHAT IS DATA VISUALISATION?

WHAT IS DATA VISUALISATION?

4



- › Present information that is intuitive and clear for the viewer
- › Turn numbers in a spreadsheet into something people can interpret and extract insights

WHY VISUALISE DATA?

5

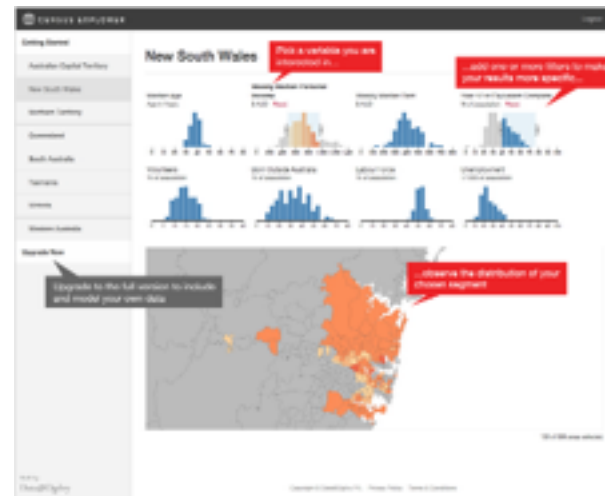
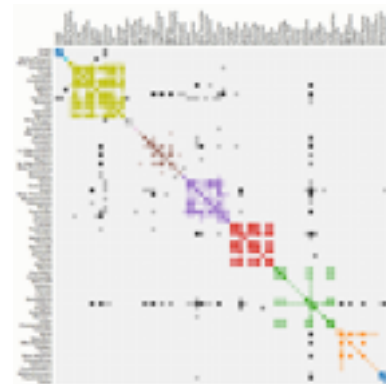
Reporting

- › Dashboards and Business Intelligence
- › Know the questions you want answers to
- › Can detect changes from the norm
- › Good for taking a 30,000 foot view of the problem



Exploring

- › Exploratory Data Analysis
- › Combines multiple data sources for single view of a problem
- › Technical analysis of data
- › Combined with modelling allows for the discovery of new problems and solutions





Easy to Use

Powerful

Advantages

- Provides a useful starting point
- Familiar to a large audience
- Prototyping and design time is reduced
- Default settings reduce the options and thinking that goes into producing a graph

- Scales to larger datasets
- Customised visualisations can create engaging visualisations
- Open-source (so free to run and extend)
- Non-obvious insights can be discovered with modelling tools
- Re-use code to produce similar charts for different data

Disadvantages

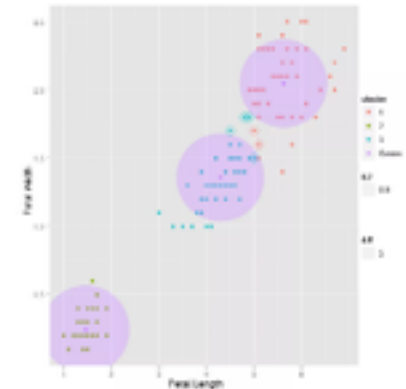
- Reproducing analysis requires lots of manual effort
- Limited to relative small data sets
- Solves known problems and cannot answer complex questions
- Licensing can be expensive

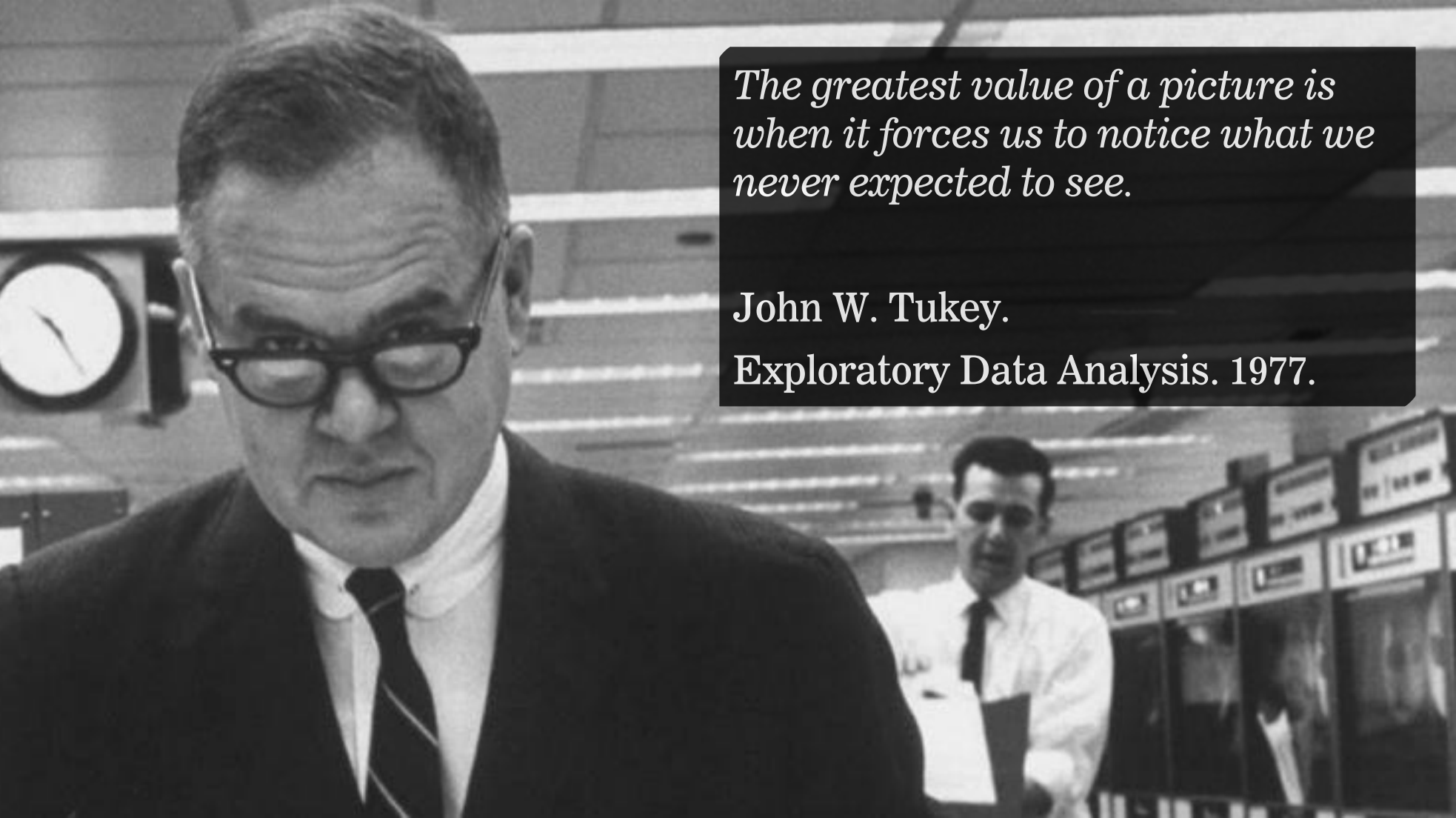
- Requires specialist skills to produce a graph
- Training and education for some of the output might be necessary

THE VALUE OF DATA VISUALISATION

9

- Communicate what's happening within the business
- Support decisions with information
- Measure and report the impact of decisions
- Discover ways to improve the business





*The greatest value of a picture is
when it forces us to notice what we
never expected to see.*

John W. Tukey.

Exploratory Data Analysis. 1977.

DATA SCIENCE PART TIME COURSE

DATA VISUALISATION LAB

DATA SCIENCE – Week 2 Day 1

DISCUSSION TIME

- **Review of last week**
- **Further Reading for Data Visualisation**
- **Check in with homework/course project**

WEEK 1 - Wednesday

Wednesday 9th Dec

☒ A Definition of DS

☒ Git-Forking

☒ Running Jupyter+Pandas

☒ Data Manipulation in R

☒ Homework

WEEK 1 - Monday

Data Science Week 1 Monday

- ☒ Course Expectations
- ☒ What the Course Involves
- ☒ Course Project Information (Basics)
- ☐ Lab-Setup
 - ☐ Anaconda Python
 - ☒ Git/Github
 - ☐ Slack

DISCUSSION TIME

Further Reading

- Edward Tufte, *The Visual Display of Quantitative Information*
- Leland Wilkinson, *The Grammar of Graphics*
- Scott Murray, *Interactive Data Visualisation for the Web* (free online)
- flowingdata.com
- New York Times (Upshot)



DATA SCIENCE - Week 2 Day 1

DISCUSSION TIME

Homework/Course Project

- **How's it going ?**
- **Purpose is to set you up so you can do a little bit of your project over the break**
- **Also, anyone interested in participating in a kaggle competition?**

kaggle