Week 10: Visualising Processes

Visual Data Analytics

University of Sydney

Outline

- Flow Charts
- PERT Chart
- Gantt charts
- Waterfall/ Funnel
- More Sankey diagrams

Motivation

- Many business processes are easier to understand if visualised.
- This is particularly the case for project management.
- In this topic we will not be visualising observations and variables as we did in other topics
- However it does have similarities to graph theory and networks.

Flow charts

Processes

- Critical to all businesses and large organisations.
- Can be complex and missing steps can be disastrous.
- Visualising processes with flow charts prevents such mistakes.
- A flow chart helps with the organisation of workflow.

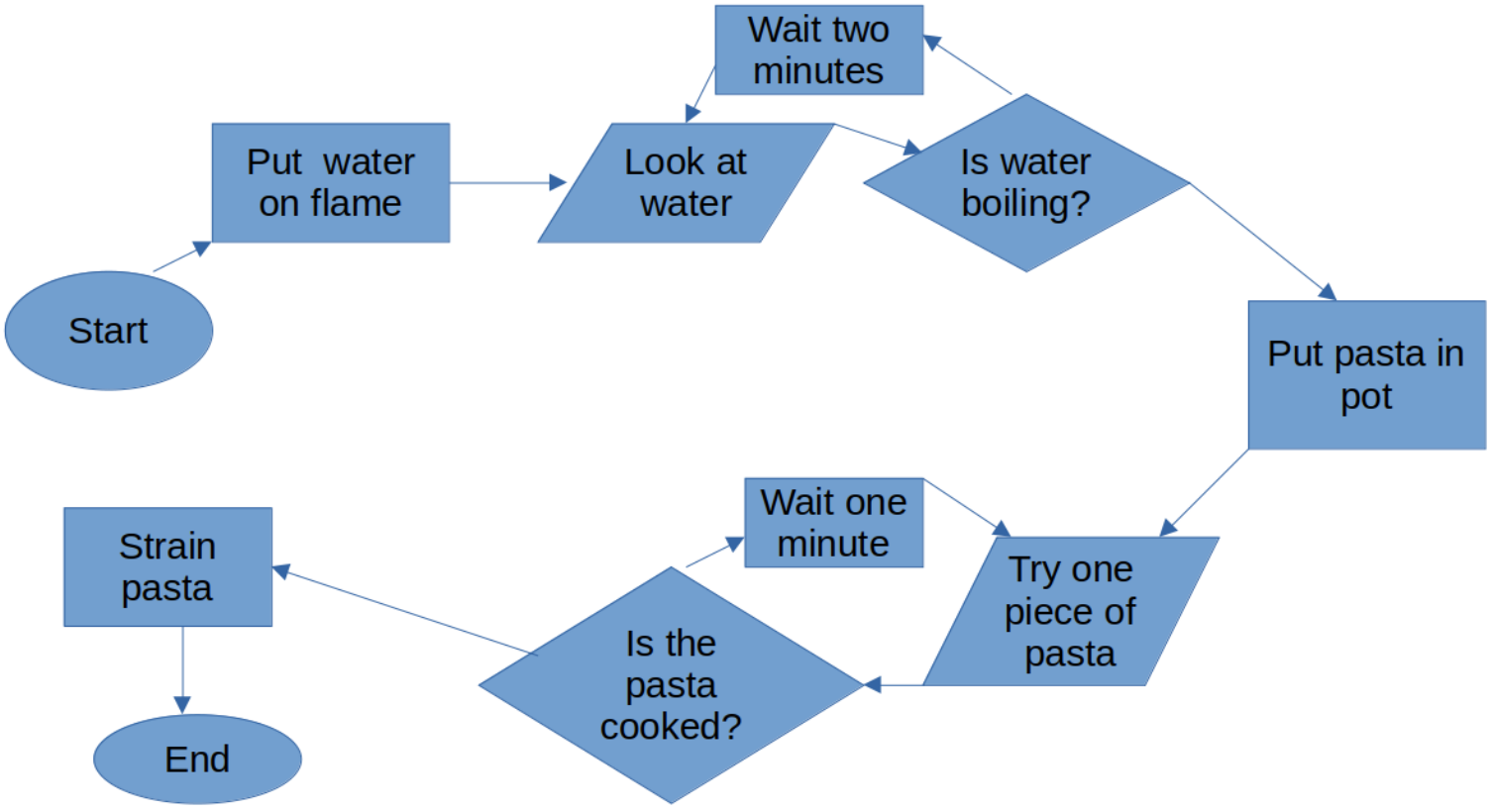
Elements of a flow chart

- Start and end
- Inputs and outputs
- Processes
- Outputs
- Connections

Flow chart symbols

Symbol	Name	Function
	Start/end	An oval represents a start or end point.
	Arrows	A line is a connector that shows relationships between the representative shapes.
	Input/Output	A parallelogram represents input or ouptut.
	Process	A rectangle represents a process.
	Decision	A diamond indicates a decision.

How to cook pasta



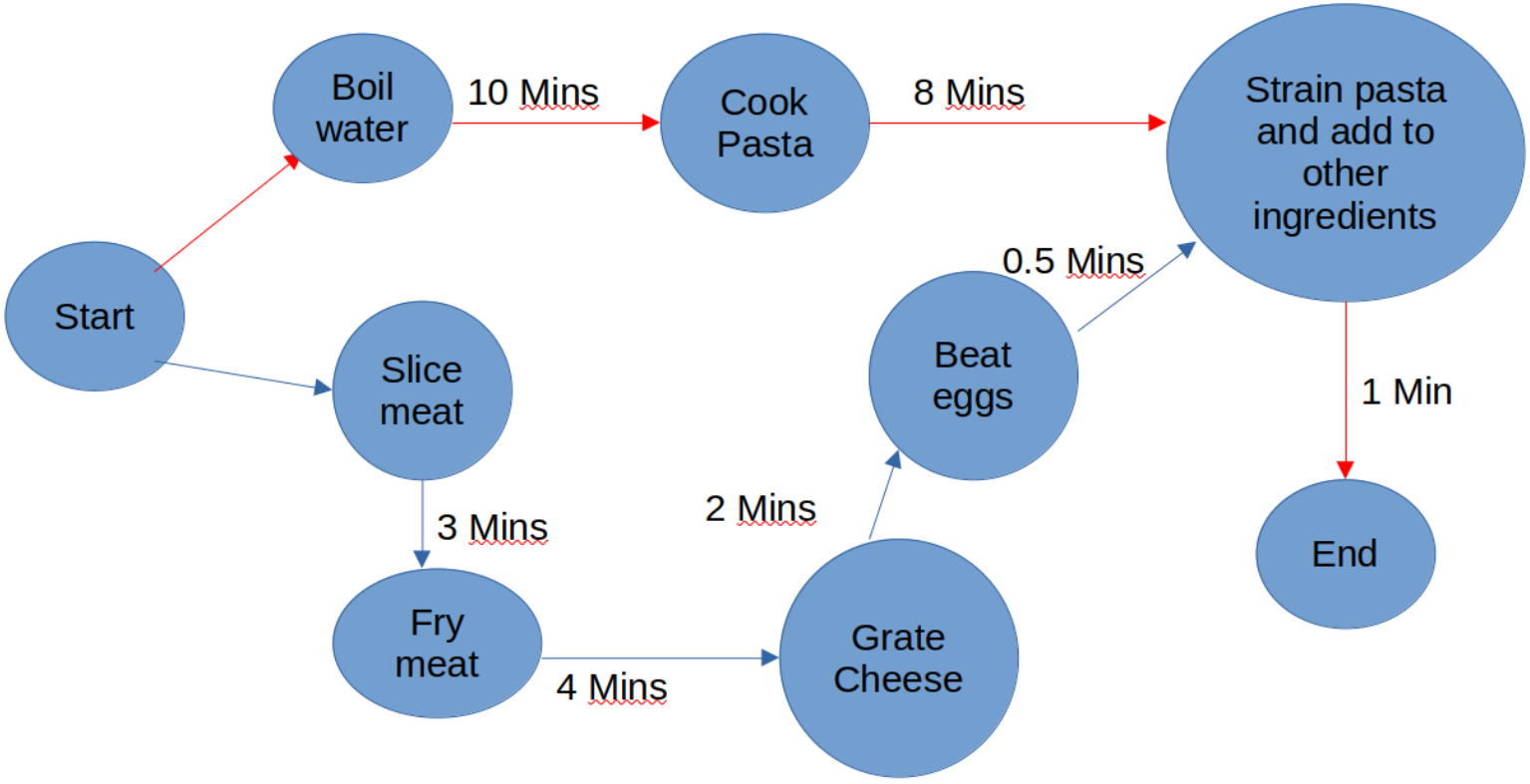
Uses of flow charts

- We can use a flow chart to summarise our own workflow as data scientists.
  - Generate questions, tidy data visualise and back to questions again.
  - Flow charts can be useful to help you organise code.
- Sometimes the best projects start out with pen and paper!

CPM/ PERT charts

- In many operations the Critical Path Method (CPM) and Program Evaluation and Review Technique (PERT) aid decision.
- The full implications of these go beyond the scope of the course.
- We will focus on the visual aspect.

More pasta


Boil water
10 Mins
Cook Pasta
8 Mins
Strain pasta and add to other ingredients
Start
0.5 Mins
Beat eggs
Slice meat
1 Min
3 Mins
2 Mins
End
Fry meat
Grate Cheese
4 Mins

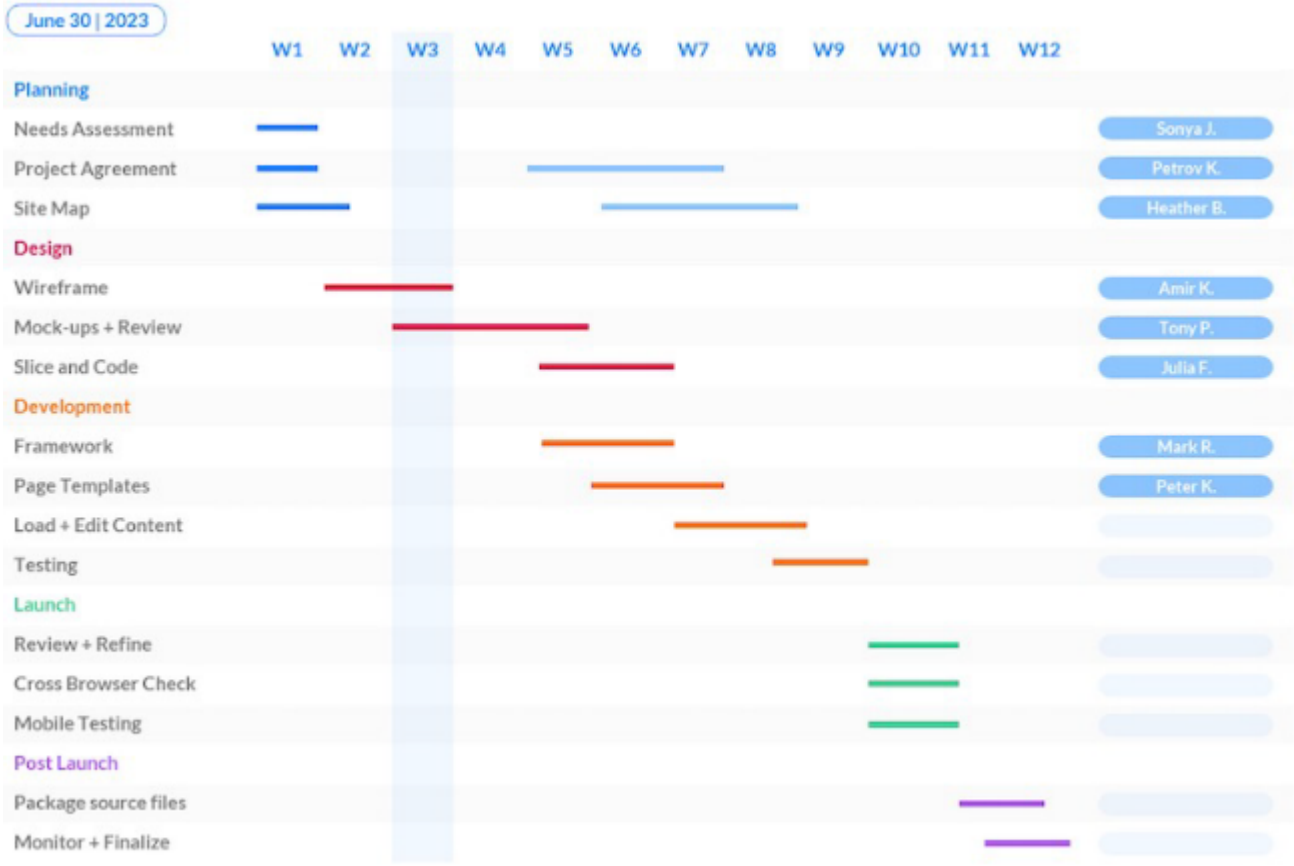
Critical path

- The purpose of CPM/PERT is to identify the *critical path*.
- In the pasta example, boiling the water and cooking the pasta take longer than the sauce.
- Any delays along the critical path leads to delays in the whole project.
- Note that we are visualising this as if it were a network.

Gantt chart

- The Gantt chart was popularised by Henry Gantt in the early twentieth century.
- It was used primarily to make industrial processes more efficient.
- It is particularly useful for showing how tasks can be done in parallel while waiting for other tasks to be completed.

Example

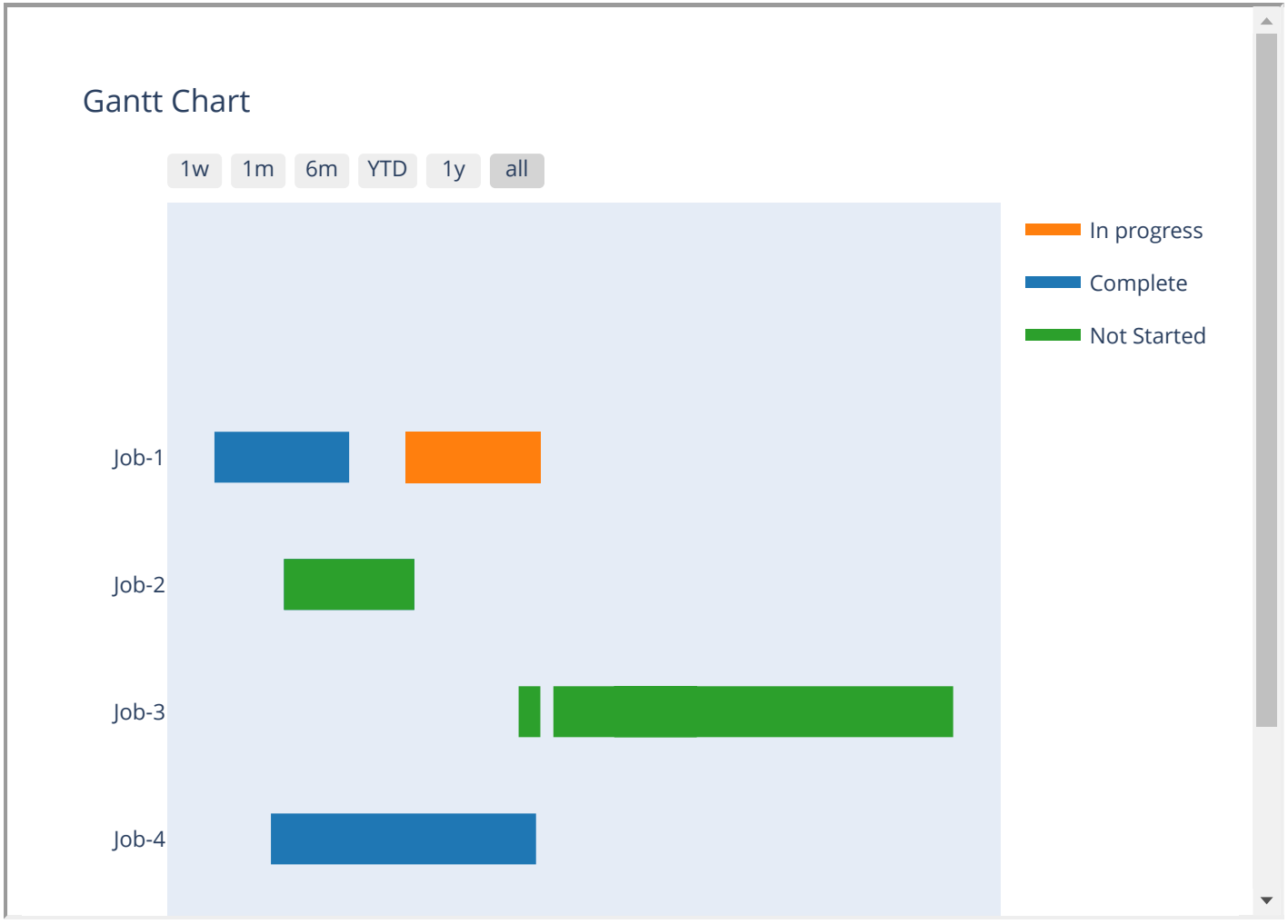

June 30 | 2023
W1
W2
W3
W4
W5
W6
W7
W8
W9
W10
W11
W12
Planning
Needs Assessment
Sonya J.
Project Agreement
Petrov K.
Site Map
Heather B.
Design
Wireframe
Amir K.
Mock-ups + Review
Tony P.
Slice and Code
Julia F.
Development
Framework
Mark R.
Page Templates
Peter K.
Load + Edit Content
Testing
Launch
Review + Refine
Cross Browser Check
Mobile Testing
Post Launch
Package source files
Monitor + Finalize

In Python

```
import plotly.figure_factory as ff

df = [dict(Task="Job-1", Start='2021-06-01', Finish='2021-07-02', Resou
      dict(Task="Job-1", Start='2021-07-15', Finish='2021-08-15', Resou
      dict(Task="Job-2", Start='2021-06-17', Finish='2021-07-17', Resou
      dict(Task="Job-2", Start='2021-06-17', Finish='2021-07-17', Resou
      dict(Task="Job-3", Start='2021-08-10', Finish='2021-08-15', Resou
      dict(Task="Job-3", Start='2021-09-01', Finish='2021-09-20', Resou
      dict(Task="Job-3", Start='2021-08-18', Finish='2021-11-18', Resou
      dict(Task="Job-4", Start='2021-06-14', Finish='2021-08-14', Resou
fig = ff.create_gantt(df, index_col='Resource',show_colorbar=True, grou
fig.write_html('gantt.html')
```

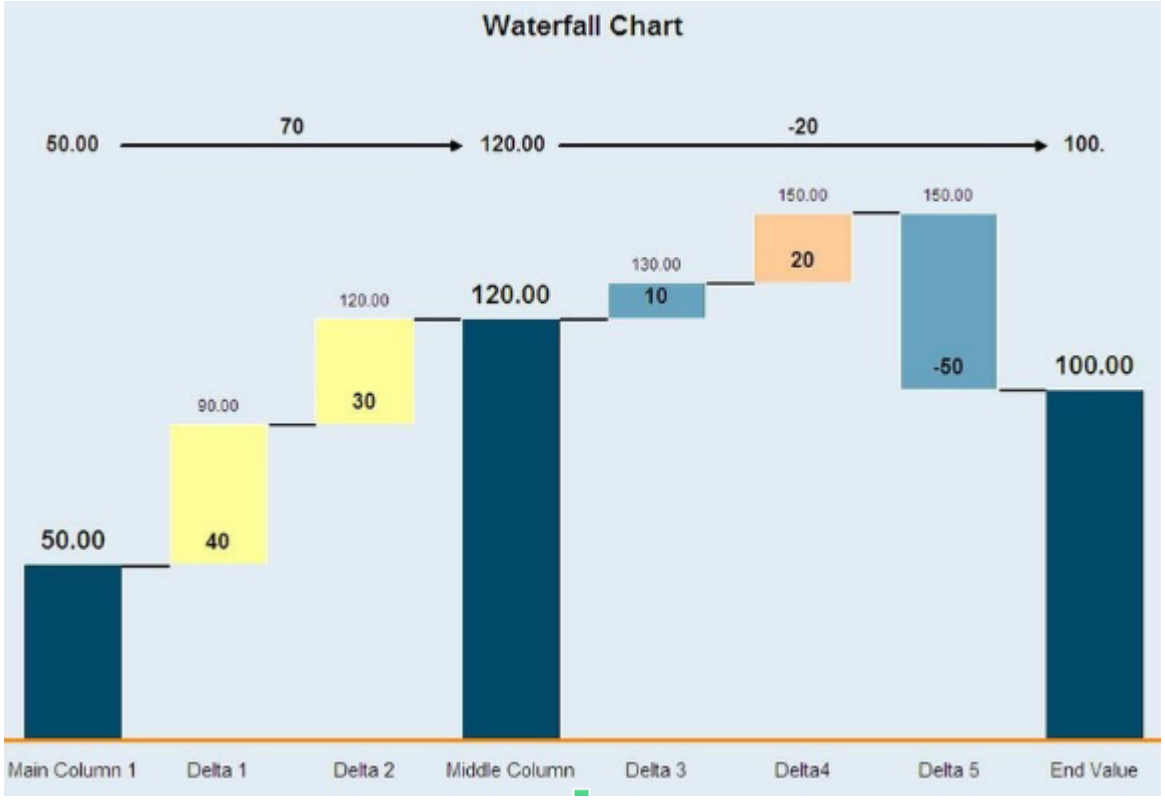
Plot


Gantt Chart
1w
1m
6m
YTD
1y
all
In progress
Complete
Not Started
Job-1
Job-2
Job-3
Job-4

Waterfall

- The waterfall chart shows time across the horizontal axis and the cumulative effects of payments on the vertical axis.
- It is useful for showing cash flows and revenue in financial analysis.
- Different colors may indicate different types of cash flow.

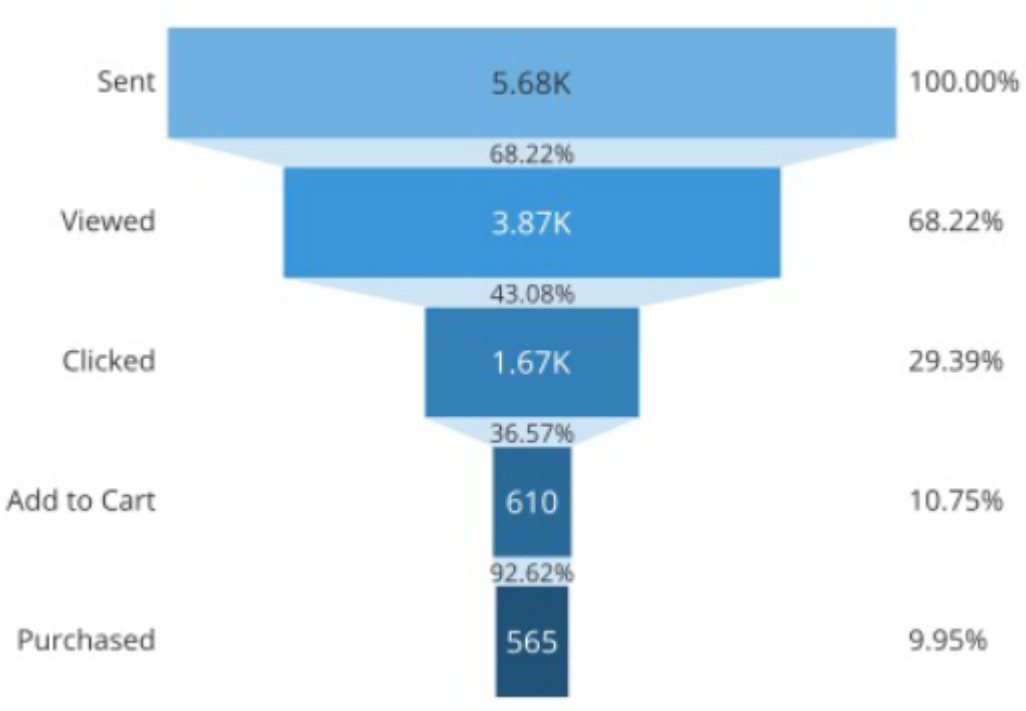
Example


Waterfall Chart
50.00
70
120.00
-20
100.
150.00
150.00
130.00
20
120.00
120.00
10
90.00
30
-50
100.00
50.00
40
Main Column 1
Delta 1
Delta 2
Middle Column
Delta 3
Delta4
Delta 5
End Value

Funnel

- The funnel shows values across multiple stages.
- Time is shown vertically
- It is a special case of a Sankey diagram.
- It is particularly useful in marketing applications

Example


Sent
5.68K
100.00%
68.22%
Viewed
3.87K
68.22%
43.08%
Clicked
1.67K
29.39%
36.57%
Add to Cart
610
10.75%
92.62%
Purchased
565
9.95%

Setup in Python

```
import plotly.express as px
data = dict(
    number=[4500, 1500, 750, 500, 250],
    stage=["Website visits", "Downloads", "Contacted by email", "Reques
fig = px.funnel(data, x='number', y='stage')
fig.write_html('funnel.html')
```

Funnel plot


stage
Website visits
Downloads
Contacted by email
Requested price
Purchased
4500
1500
750
500
250

Python

- Python code was not shown for every plot in this lecture.
- While these plots are popular in business
  - They require dedicated packages in Python or
  - They can be customised with the usual visualisation libraries.
- Often they are easier to create with other software.

Wrap-up

Conclusions

- There are many ways to visualise processes.
  - Often these rely on a graph / network visualisation.
  - In other cases they are quite unique.
- While these are often used for big organisations, you can use them as individuals to plan out your own workflows.

Questions