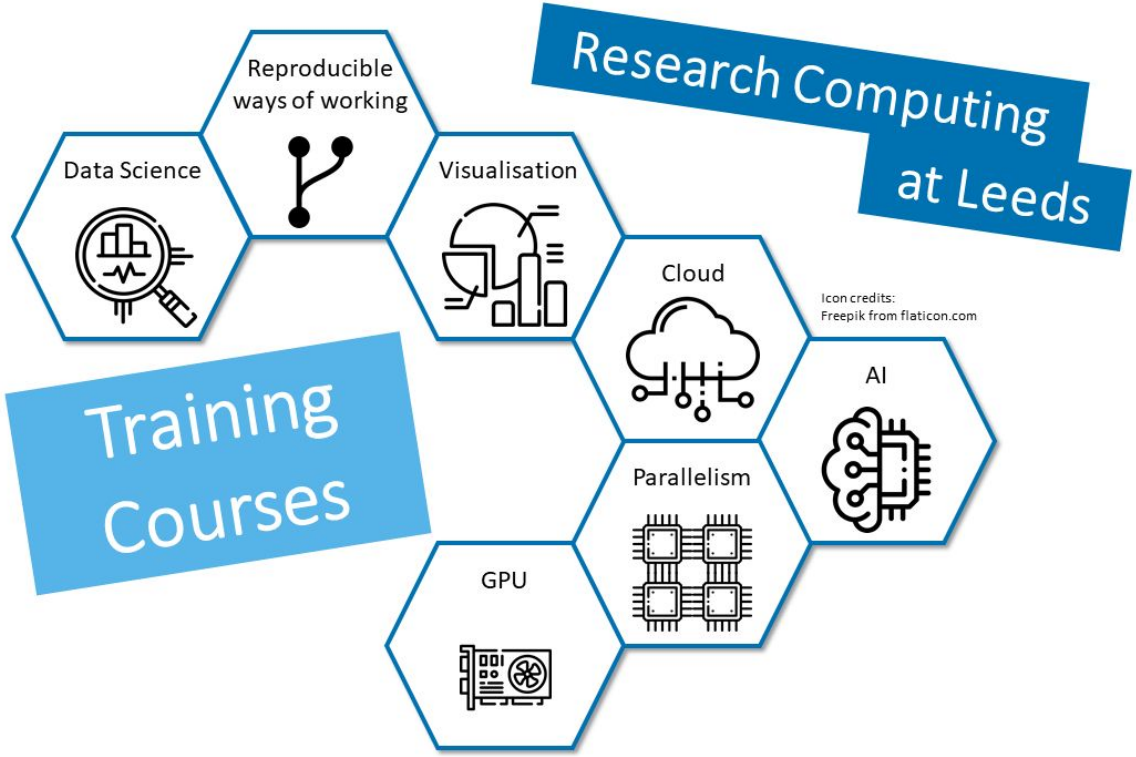
SWD7: Introduction to Data Visualisation with Python



Research IT Website | Research IT Query: bit.ly/arc-help | Course Materials Homepage

Research Computing at Leeds

- Here to support research and researchers:
  - Provide training
  - Support users of Grid and Cloud Computing platforms
  - Provide consultancy:
    - To develop project proposals
    - To help recruit people with specialist skills
    - To collaborate directly on research projects
- For details, please see our website
- Contact us via the IT Service Desk

What is data visualisation?

"Data visualisation is the graphical representation of information and data.

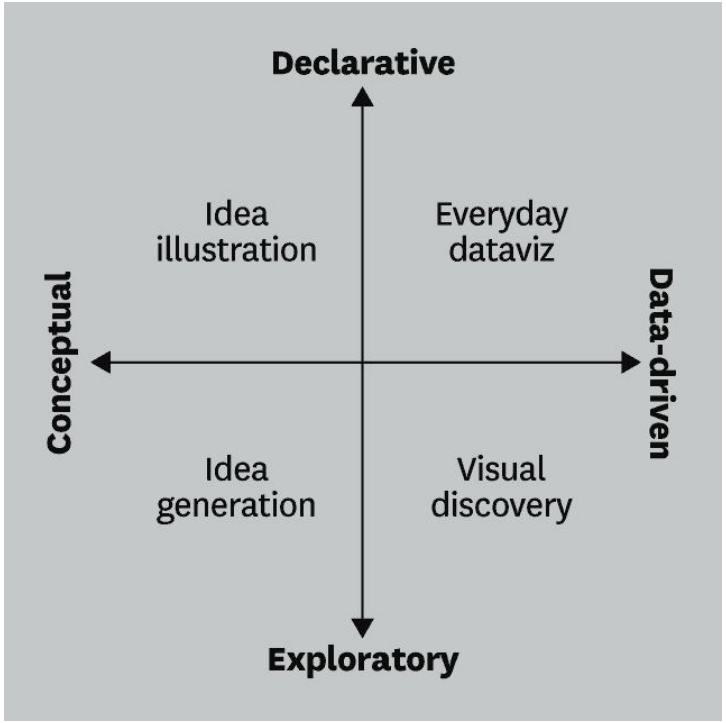
By using visual elements like charts, graphs and maps, data visualisation tools provide an accessible way to see and understand trends, outliers and patterns in data."

Tableau Documentation website

What is data visualisation?

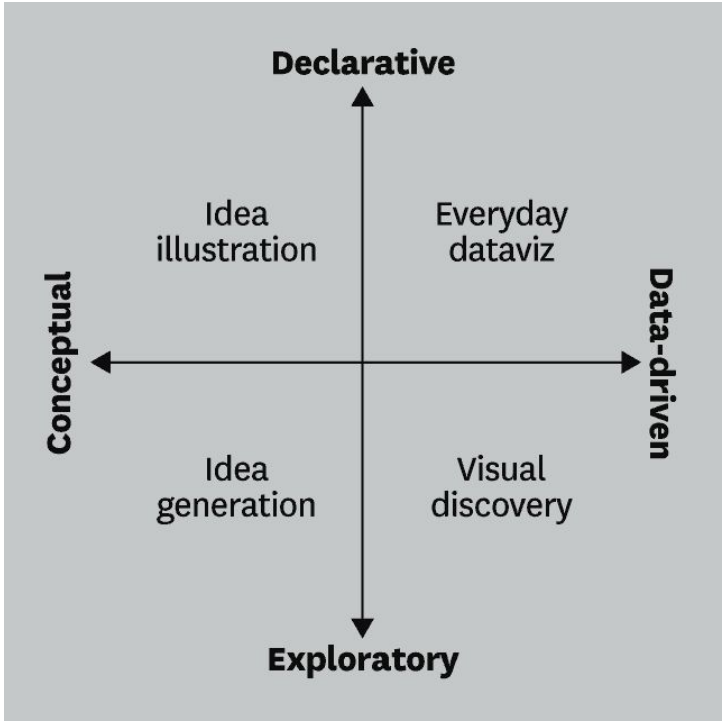
According to Harvard Business Review, successful data visualisation requires you to ask two questions before beginning:

- Is the information **conceptual** or **data-driven**?
  - We'll be mainly focusing on data-driven, *quantitative* data vis in this course
- Am I **declaring** something or **exploring** something?
  - We will be looking at both approaches during this course!

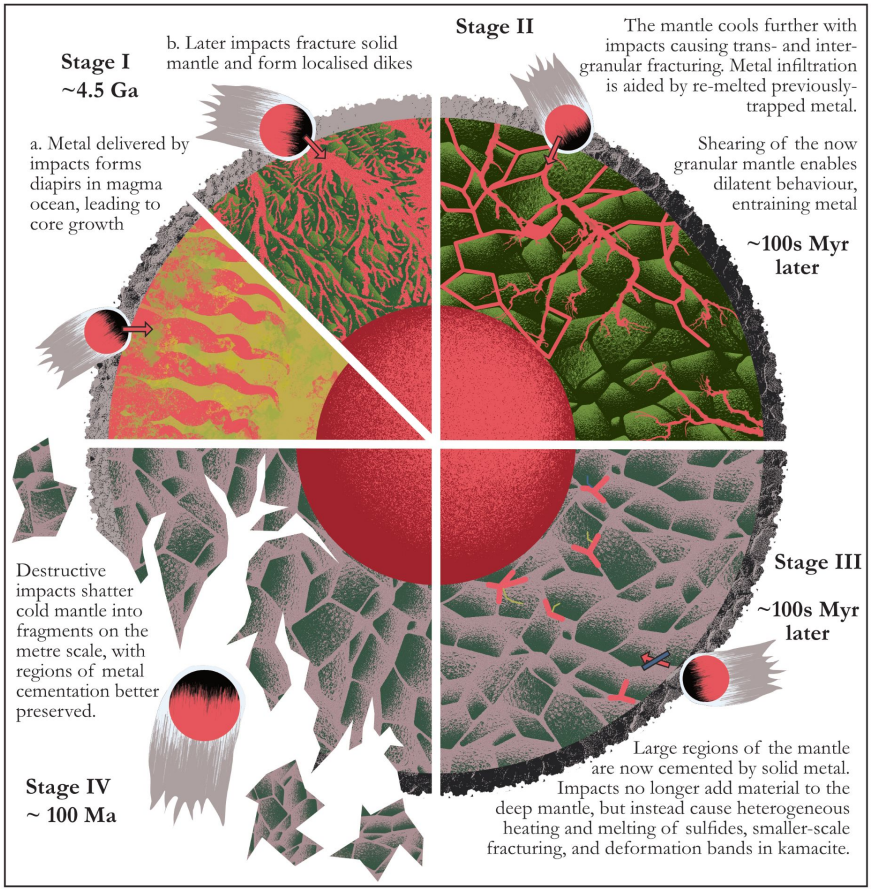


Harvard Business Review

Diagrams, flow-charts, box-diagrams, Cartoons, overviews



[Harvard Business Review]

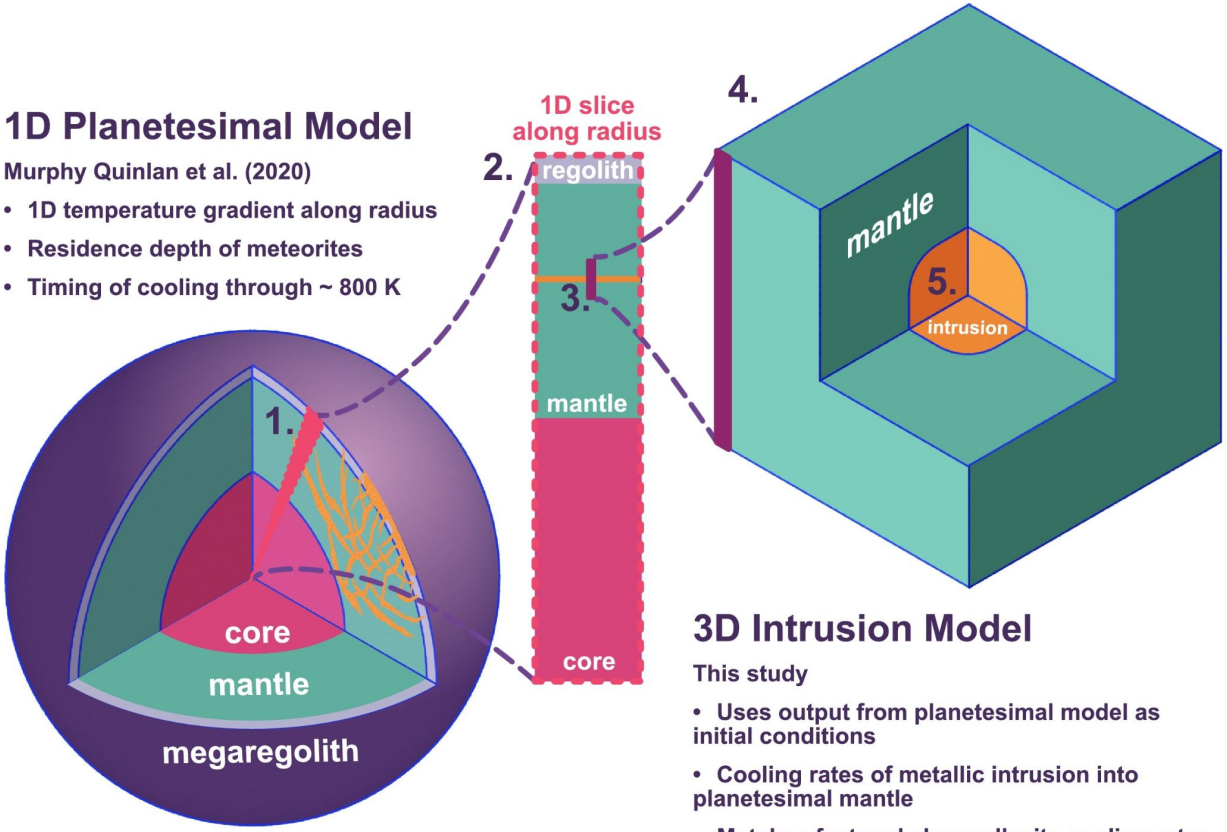


© Maeve Murphy Quinlan, 2024

1D Planetesimal Model

Murphy Quinlan et al. (2020)

- 1D temperature gradient along radius
- Residence depth of meteorites
- Timing of cooling through ~ 800 K

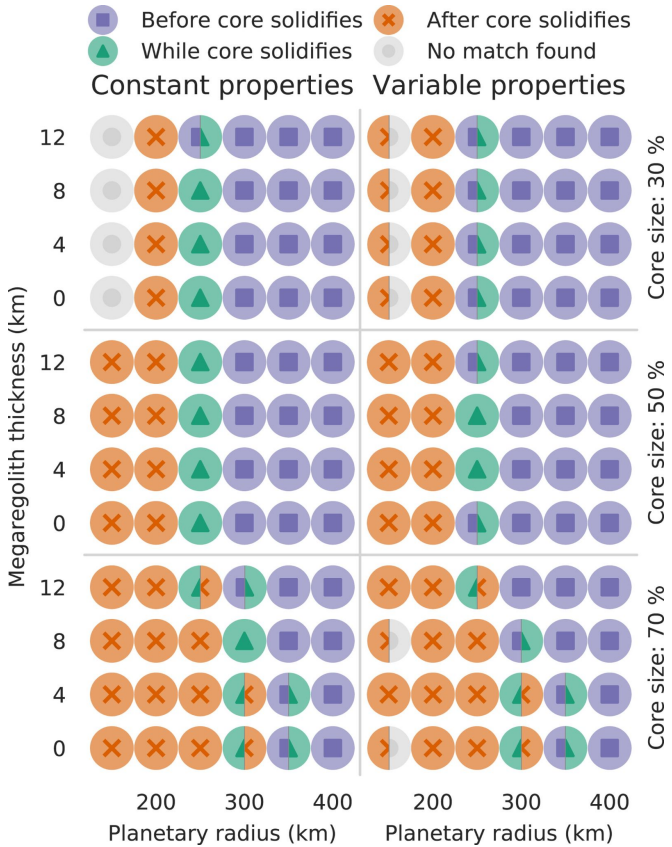


3D Intrusion Model

This study

- Uses output from planetesimal model as initial conditions
- Cooling rates of metallic intrusion into planetesimal mantle
- Matches fast and slow pallasite cooling rates

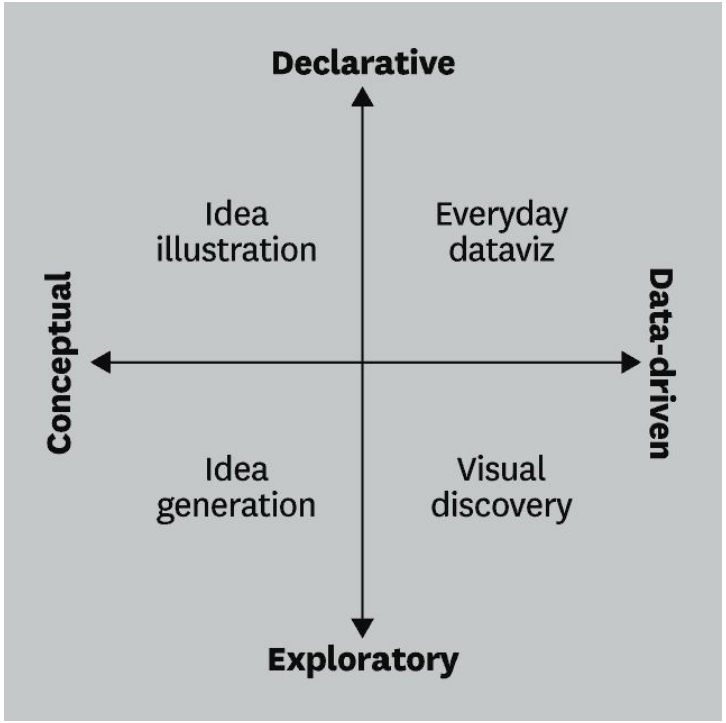
Murphy Quinlan et al., 2023



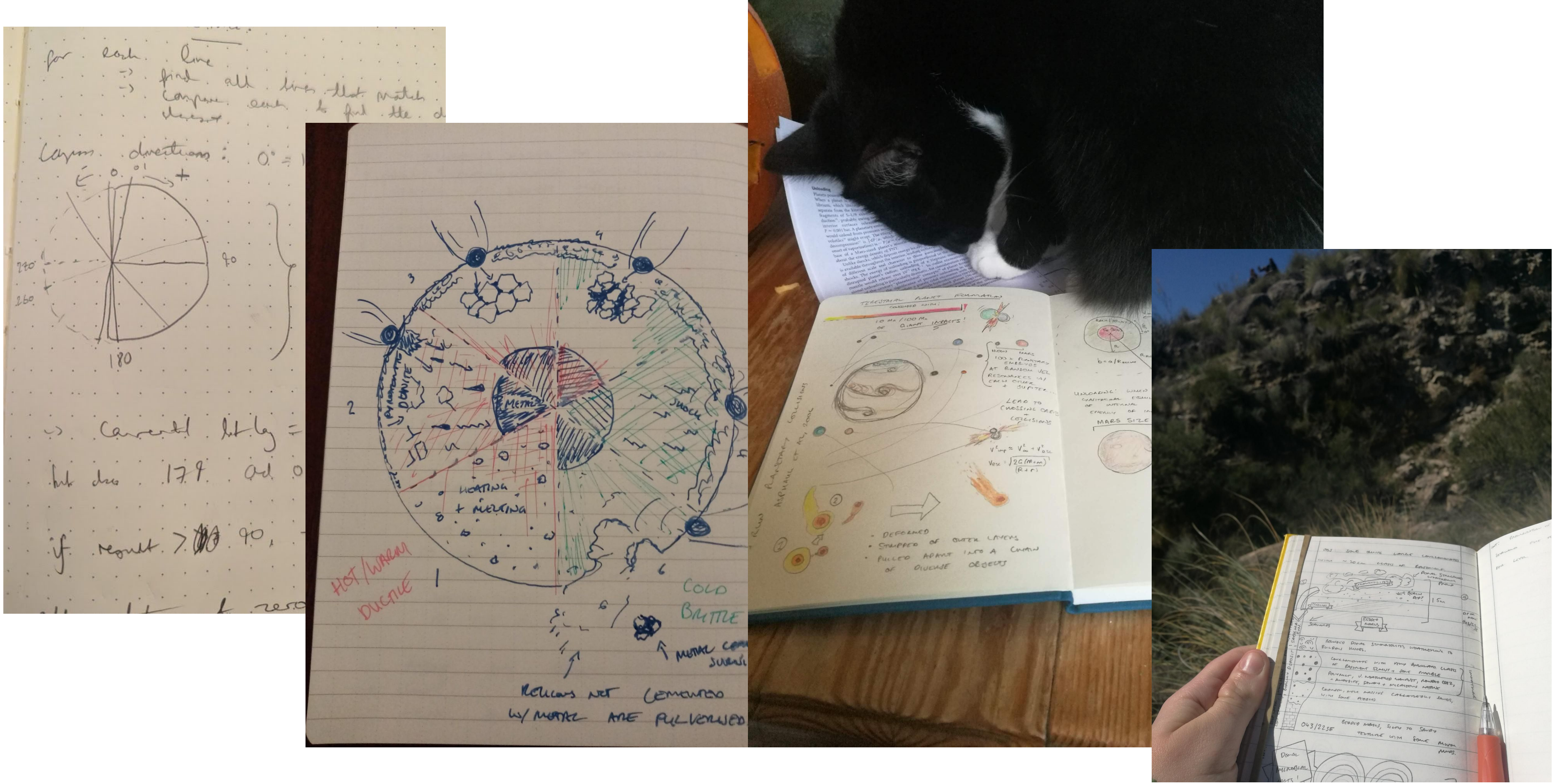
Murphy Quinlan et al., 2021

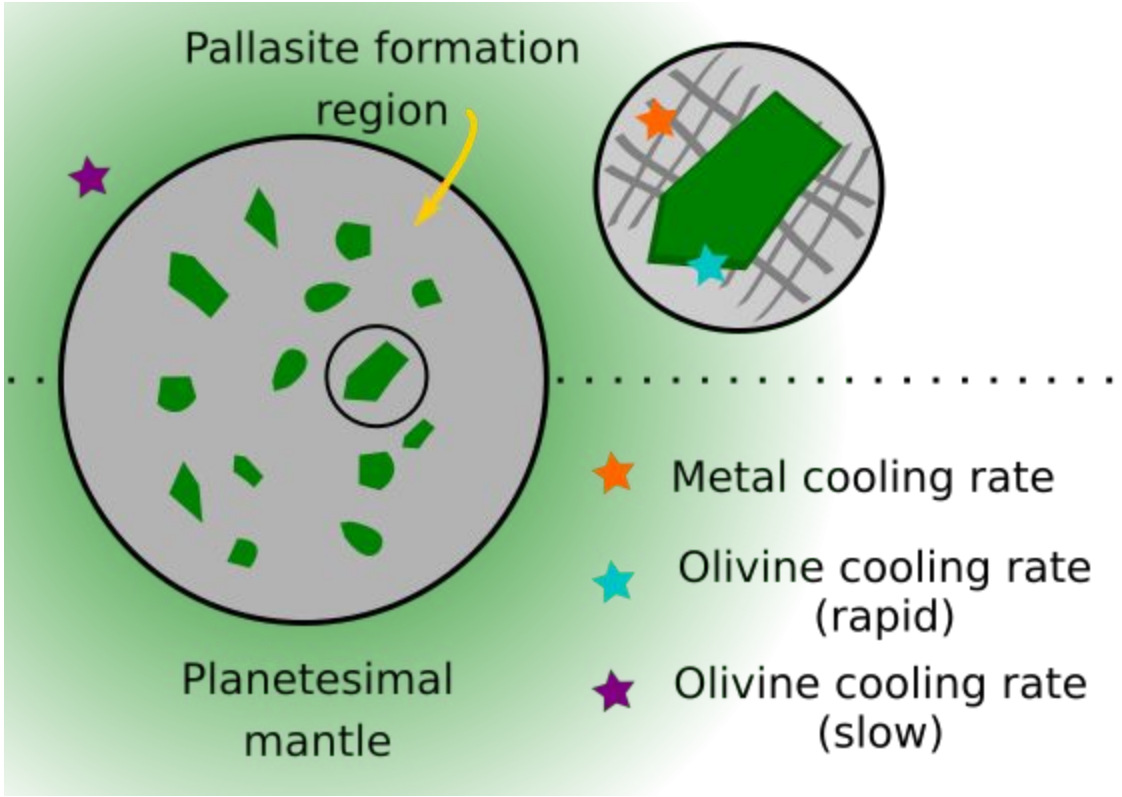
Diagrams, flow-charts, box-diagrams, Cartoons, overviews

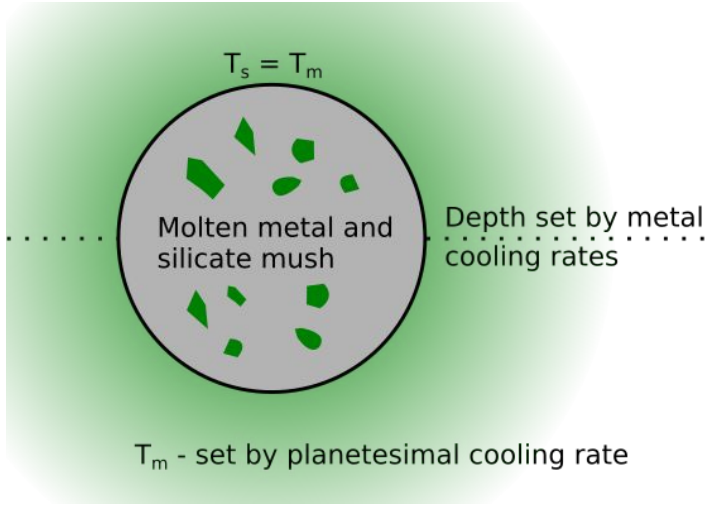
Sketched notes in your field notebook, supervisor meeting notes, doodles



Harvard Business Review

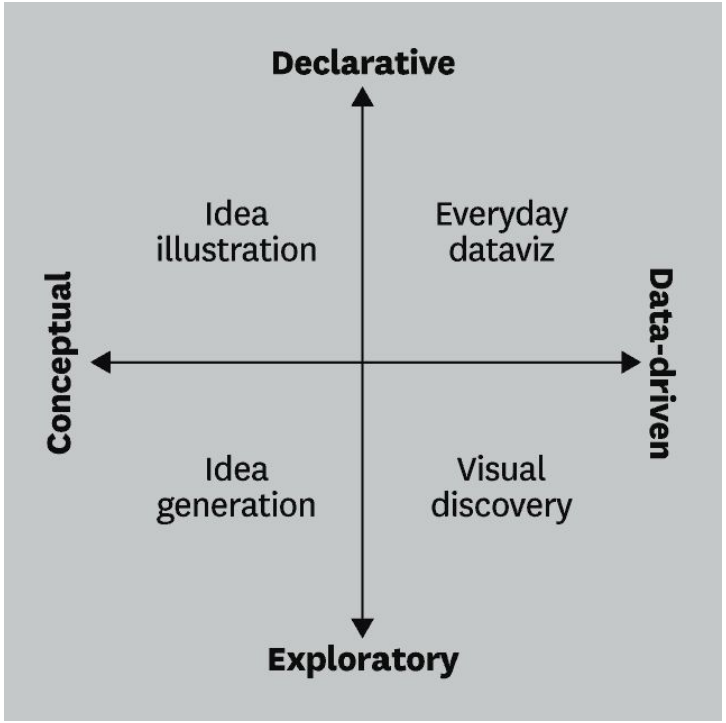

HOT/WARM
DUCTILE
COLD
BRITTLE
HEATING
+ MELTING
180


Pallasite formation region
Planetesimal mantle
Metal cooling rate
Olivine cooling rate (rapid)
Olivine cooling rate (slow)


$T_s = T_m$
Molten metal and silicate mush
Depth set by metal cooling rates
T_m - set by planetesimal cooling rate

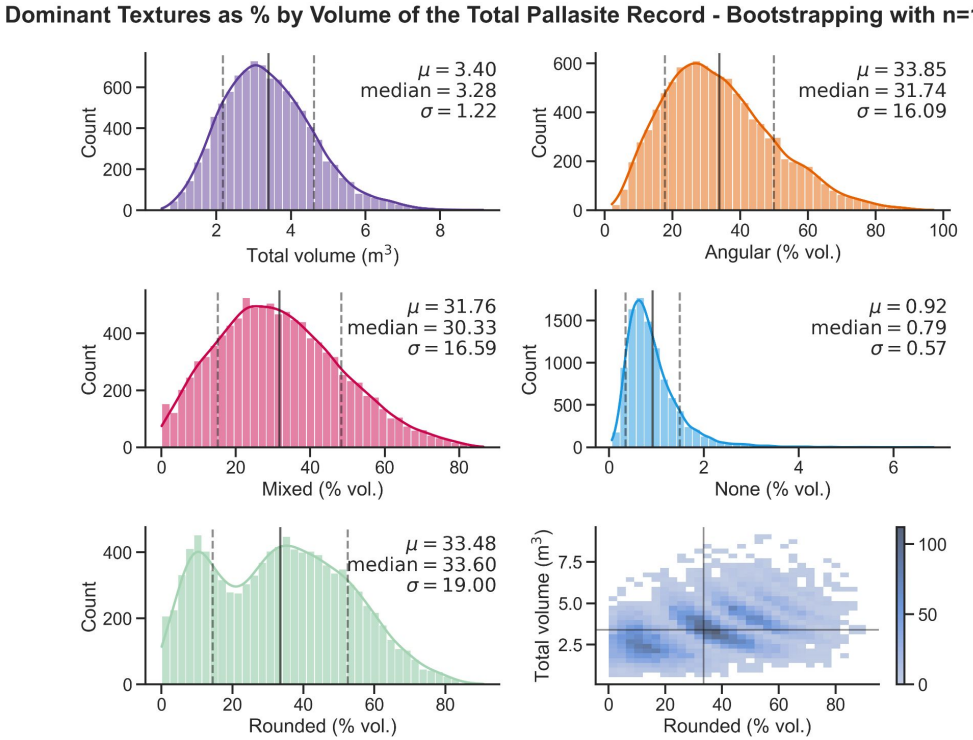
Diagrams, flow-charts, box-diagrams, Cartoons, overviews

Sketched notes in your field notebook, supervisor meeting notes, doodles

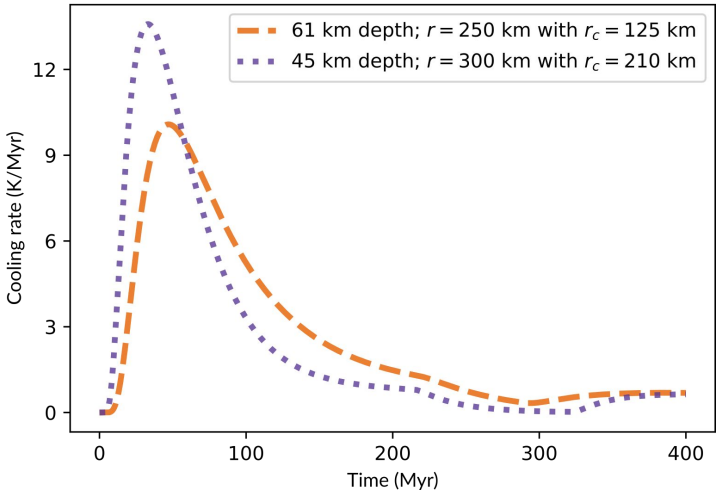


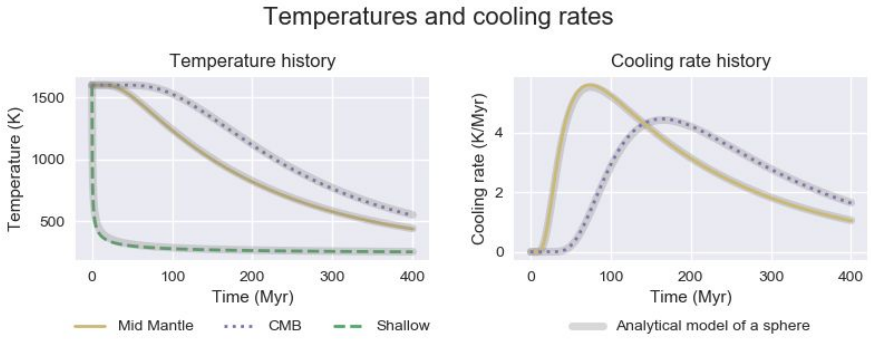
Scatter plots, line plots, interactive figures, colour-coded tables

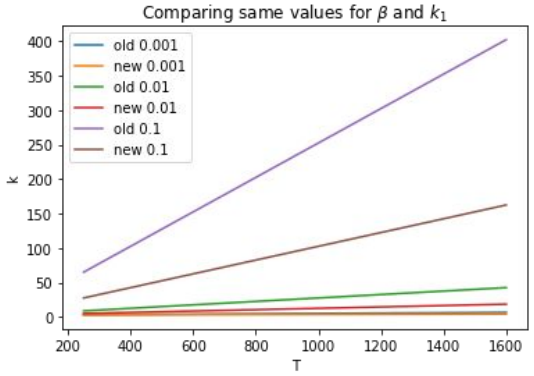
[Harvard Business Review]

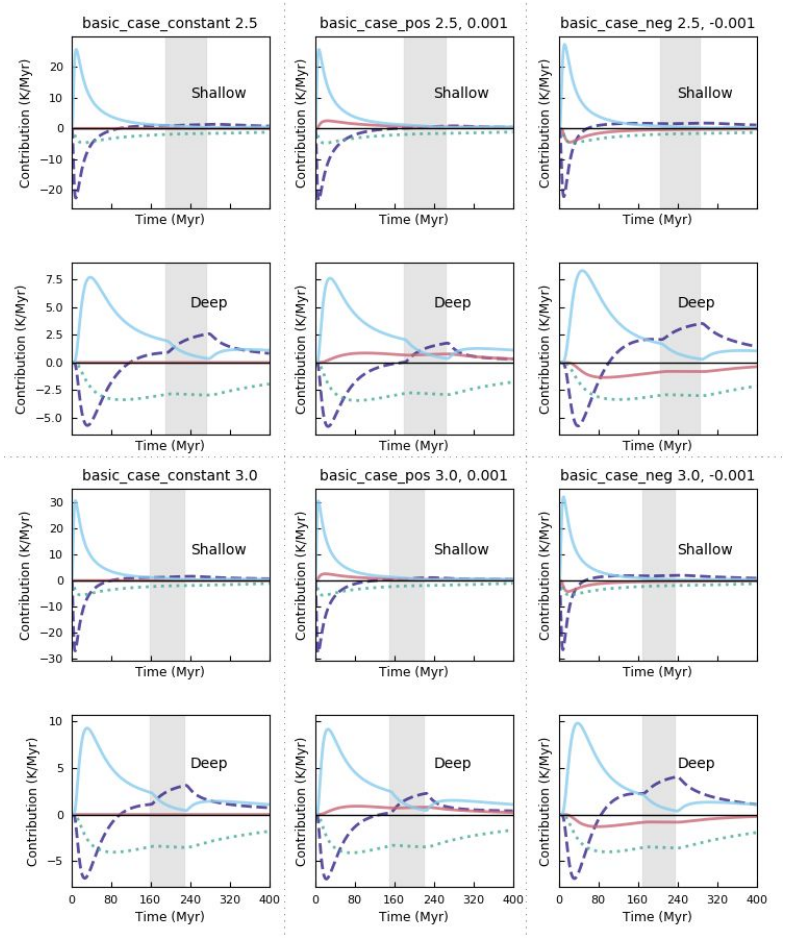


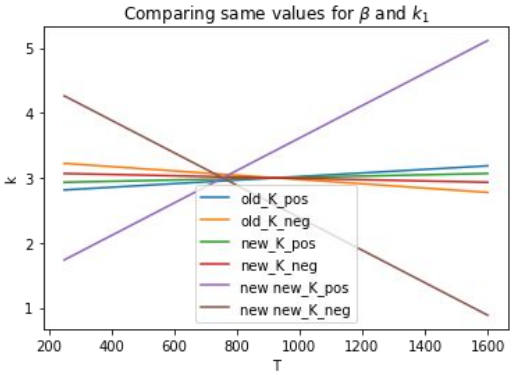
Cooling rates at depth of pallasite genesis





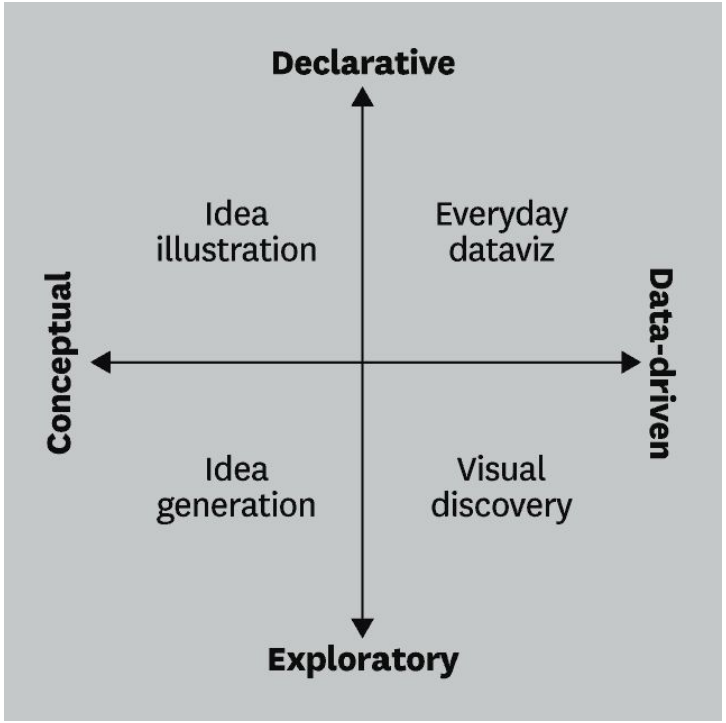

Comparing same values for β and k_1
old 0.001
new 0.001
old 0.01
new 0.01
old 0.1
new 0.1
k
T


basic_case_constant 2.5
basic_case_pos 2.5, 0.001
basic_case_neg 2.5, -0.001
basic_case_constant 3.0
basic_case_pos 3.0, 0.001
basic_case_neg 3.0, -0.001
Shallow
Deep
Contribution (K/Myr)
Time (Myr)


Comparing same values for β and k_1
old_K_pos
old_K_neg
new_K_pos
new_K_neg
new new_K_pos
new new_K_neg
k
T

Diagrams, flow-charts, box-diagrams, Cartoons, overviews

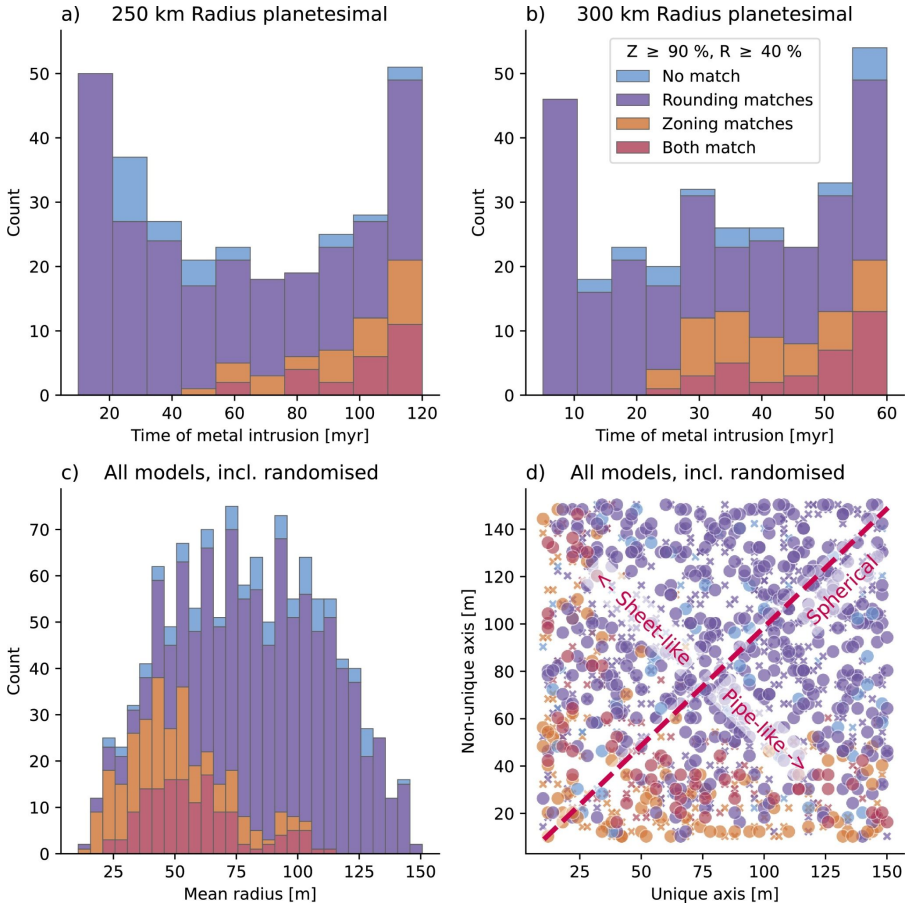
Sketched notes in your field notebook, supervisor meeting notes, doodles



Static well-designed (hopefully) plots, polished dashboards

Scatter plots, line plots, interactive figures, colour-coded tables

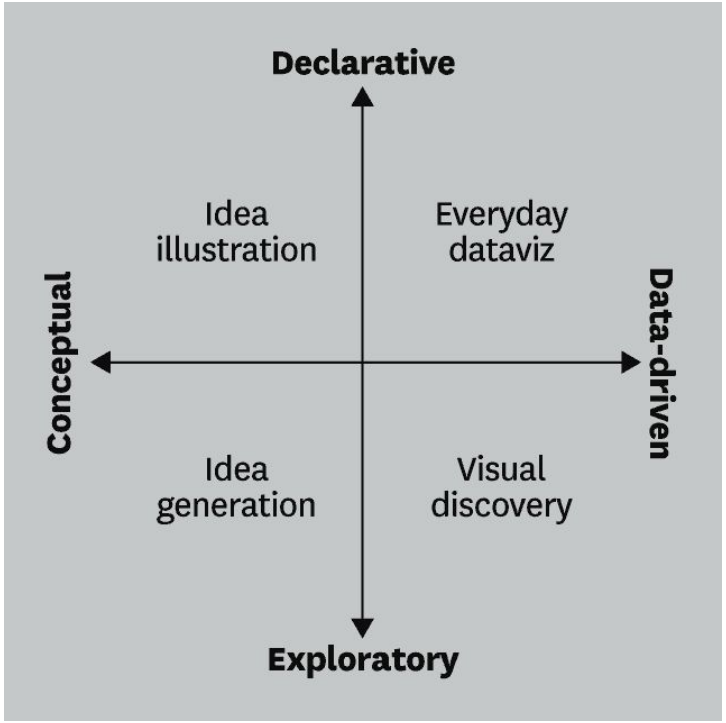
Harvard Business Review



Murphy Quinlan et al., 2023

Diagrams, flow-charts, box-diagrams, Cartoons, overviews

Sketched notes in your field notebook, supervisor meeting notes, doodles



Static well-designed plots, polished dashboards

Scatter plots, line plots, interactive figures, colour-coded tables

Harvard Business Review

Good data visualisation


Terribl
e
Good data
visualisation

What constitutes terrible data vis?

What's needed for Good Data Visualisation?

1. Evidence-based best practises for **objective** and **truthful** data representation

2. **Technical skills** in tools to build graphics

3. **Domain specific knowledge** of the underlying data set

4. Creativity, **design** principles, taste

What's needed for Good Data Visualisation?

1. Evidence-based best practises for **objective and truthful** data representation

2. **Technical skills** in tools to build graphics

3. **Domain specific knowledge** of the underlying data set

4. Creativity, **design** principles, taste

Why Python as a tool?

- Flexible, human readable programming language
- Commonly used for:
  - Data analytics
  - Scientific programming
  - Data visualisation
- Huge collection of specialised libraries to solve specific research problems
- Open Source

Why Python as a tool?

- Allows us to create reproducible graphics and share the source code
- Enables us to build efficient workflows and research pipelines, with data cleaning, statistical analysis and plotting in the same language
- Many different library for various research domain-specific plots, data types, and visualisations

But other tools also useful...

- Python is not well suited to building more infographic or illustrative charts
  - It is possible, and it is still very useful as part of the workflow
- Sometimes better suited custom programmes exist for the specific problem, for example QGIS for mapping
  - Again, there can be use for Python-assisted workflows

If you are building a data-intensive plot, or want to be able to reproduce the plot easily for updated data or a different dataset, a scripted option (Python) is a better choice than an illustration package

What's needed for Good Data Visualisation?

1. Evidence-based best practises for **objective and truthful** data representation

2. **Technical skills** in tools to build graphics

3. **Domain specific knowledge** of the underlying data set

4. Creativity, **design** principles, taste

What's needed for Good Data Visualisation?

1. Evidence-based best practises for **objective and truthful** data representation

2. **Technical skills** in tools to build graphics

What's needed for Good Data Visualisation?

Theory section:

1. Evidence-based best practises for **objective and truthful** data representation

Practical section:

2. **Technical skills** in tools to build graphics

Introductions poll