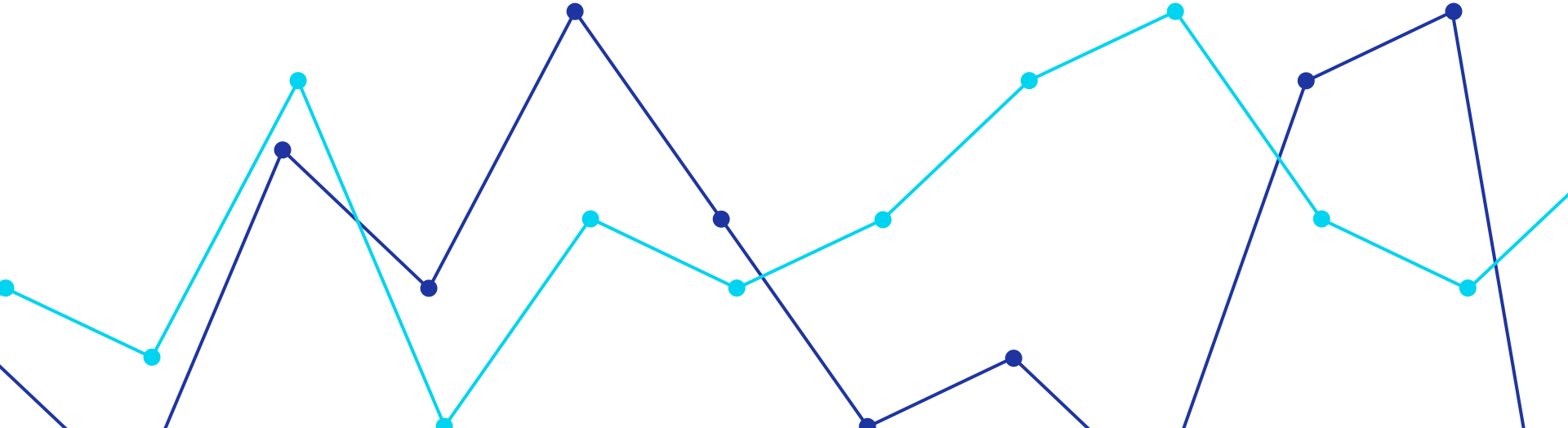
Recommender Systems Challenge

Arianna Galzerano, Francesco Fulco Gonzales



Project Timeline

Data exploration

Analysis of URM and ICM data

Models exploration and evaluation

Hyperparameters optimization and base models comparison

Hybrids

Implementation of different hybrid models starting from the best base models

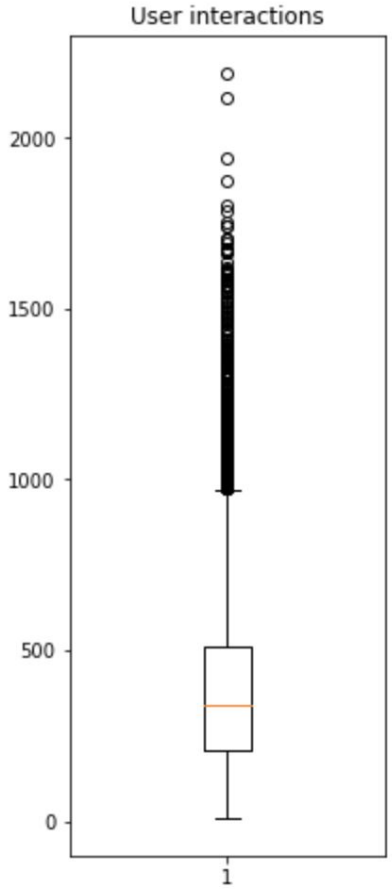
Final model

Structure of the final hybrid model

Cross-validation & other techniques

Techniques to improve performance and avoid overfitting

Understanding the data



- Clean dataset: no missing values or improper data
- All users have at least tens of interactions: CBF not needed

URM

- 5.29M Interactions
- 13650 Users
- >10 I4U

ICM

- 7 genres
- 12 subgenres
- 112 channels
- 358K events

Users

N. of interactions in URM:
min= 9, max 2191

ICMS:

Genre, subgenre, channels, event

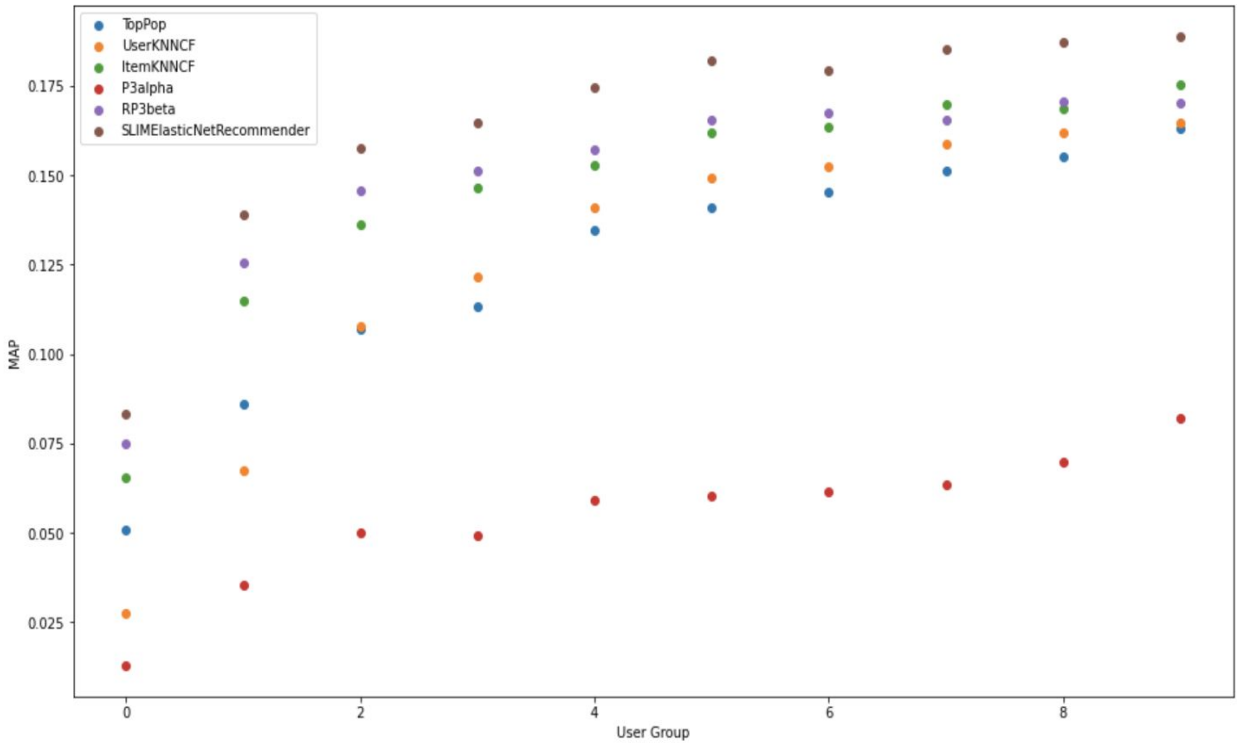
Base Models Evaluation

Base Model	***Valid MAP@10***
SLIMElasticNet	0.2391579
EASE-R	0.2393846
IALS implicit	0.2324677
PureSVD	0.2309224
RP3beta	0.2250362
ItemKNNCF	0.2131692
SLIMBPR	0.2033459

Some techniques

- Stacking URM - ICM
- Train validation split: 80:20 (initially 90:10)
- Bayesian Optimization from course repo
- Change of parameters range for different models during optimization and search
- Training both locally and in the cloud

Preprocessing



User Grouping

We tried segmentation into 2 to 10 groups of users based on n. Of interactions

Stacking

For some of the base models it was useful for performance improvement (EASE-R and Rp3beta)

Hybrids techniques and implementation

Linear Combination

Weighted sum of scores
HYBRID RATINGS class

Similarity Merge

Weighted sum of similarity matrices
HYBRID SIMILARITY CLASS

List Combination

Combination of the recommendation lists of the submodels of the hybrid

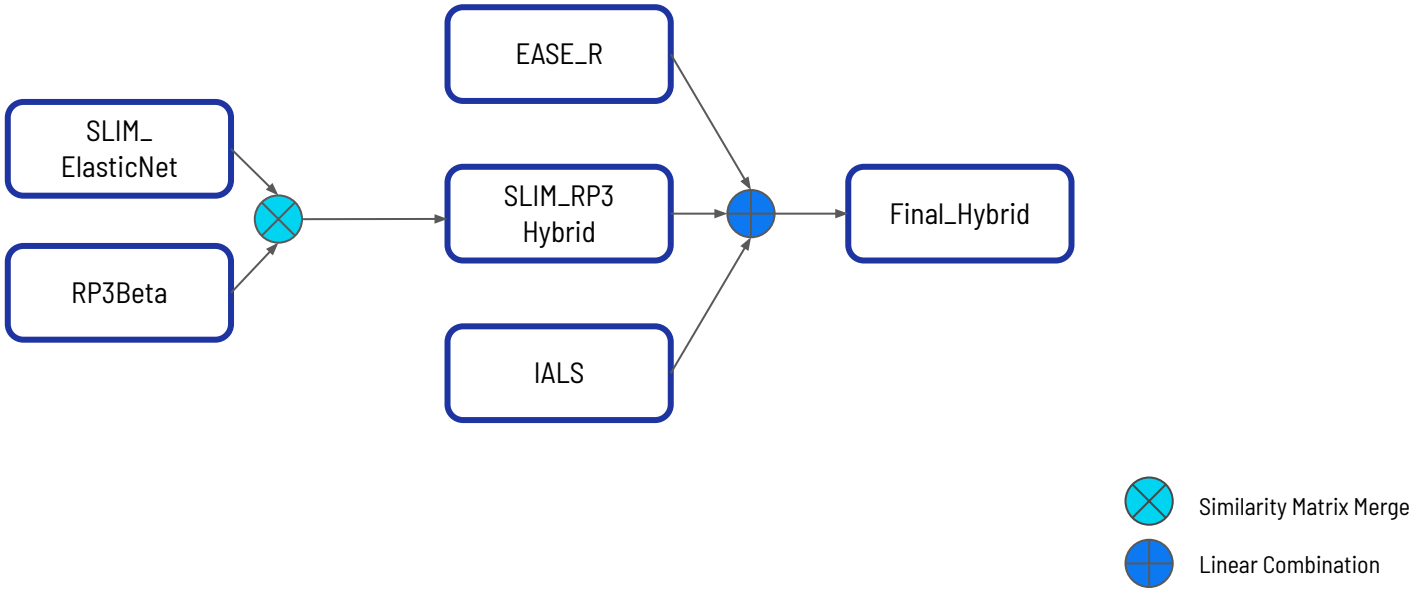
Cotraining for optimization

Joint hyperparameter optimization of models

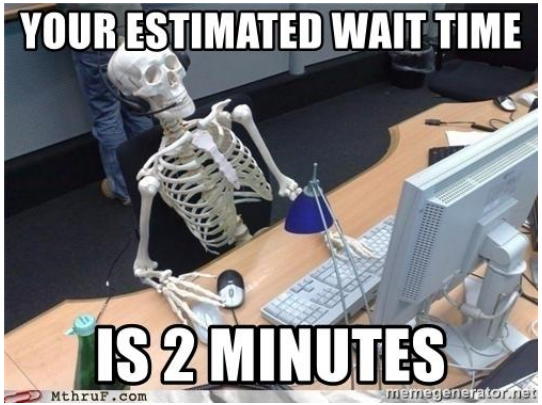
Private leaderboard MAP score: **0.48656** (7th place)

Final Hybrid Model

Stacking URM with:
- ICM channel for SLIM
- ICM event for EASE_R



Other techniques



k-fold cross validation

- k=5 to keep 80:20 split
- Bayesian hyperparameter tuning on cross-validated MAP
- Robustness check for private leaderboard

Selective Cotraining

- Fixed some of the most compute-intensive models to finetune the others

IALS from implicit library

- Significant speed up in training time
- Leveraging Cython and multithreading


Thank you for the attention!
Any questions?
Code and optimization results available on github:
@fulcus & @arigalzi

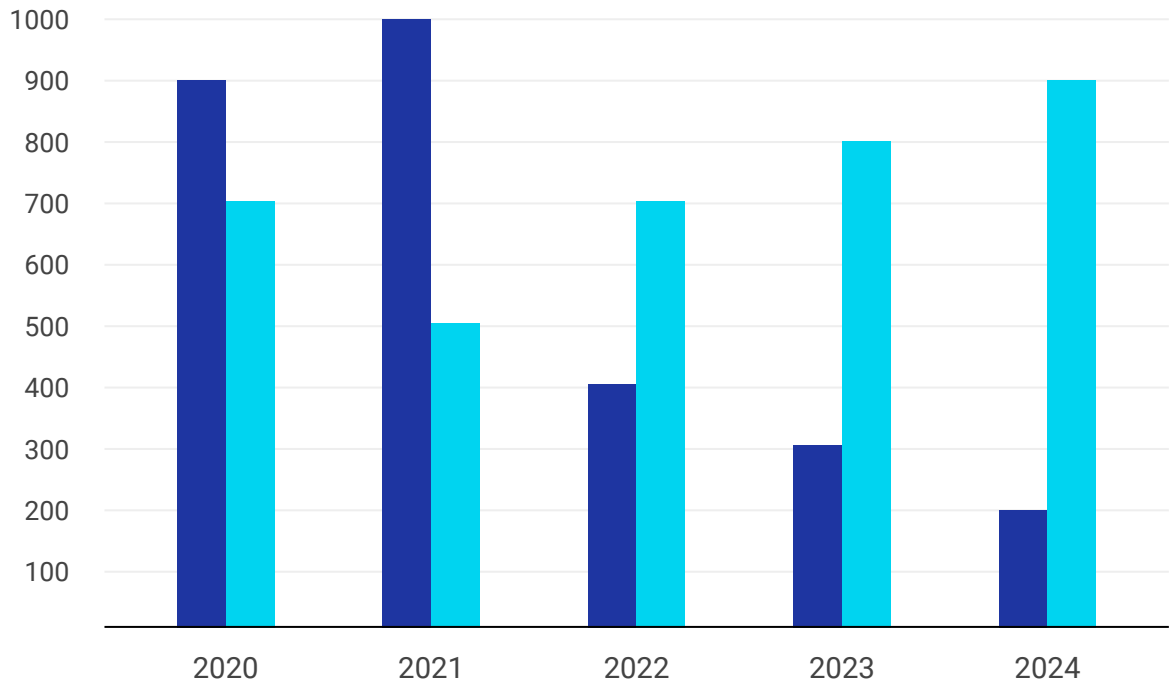
Data Charts Infographics

Product 1

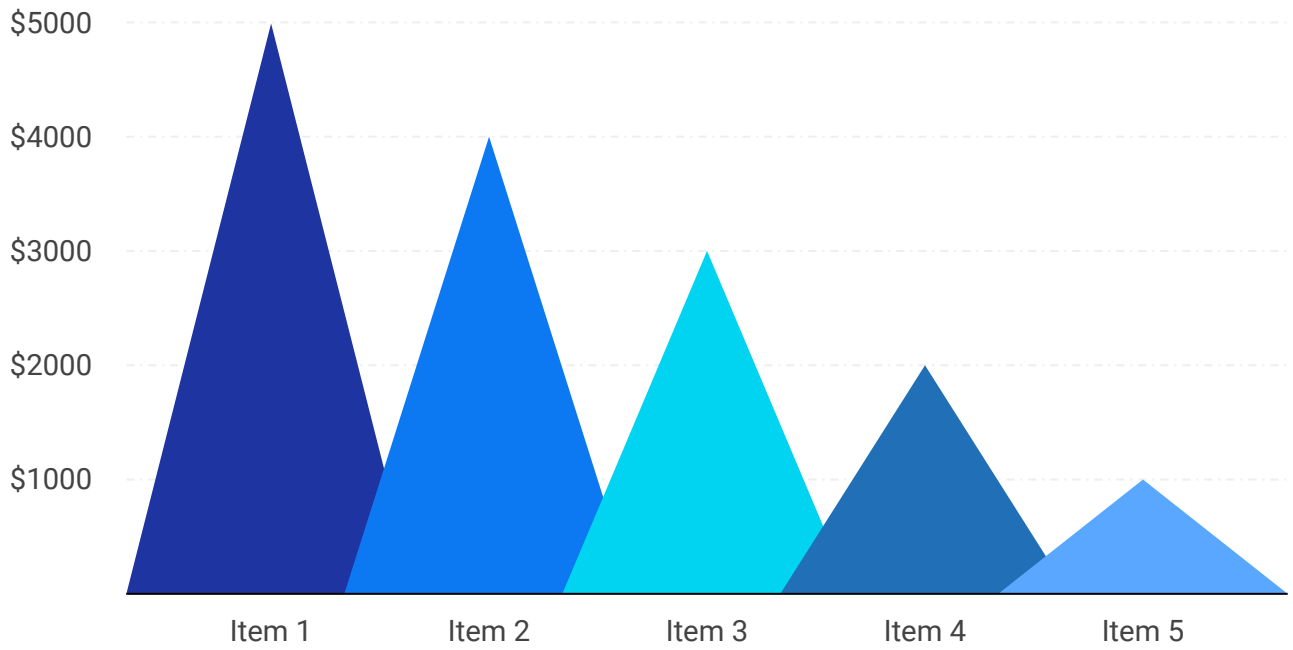
Venus is the second planet from the Sun

Product 2

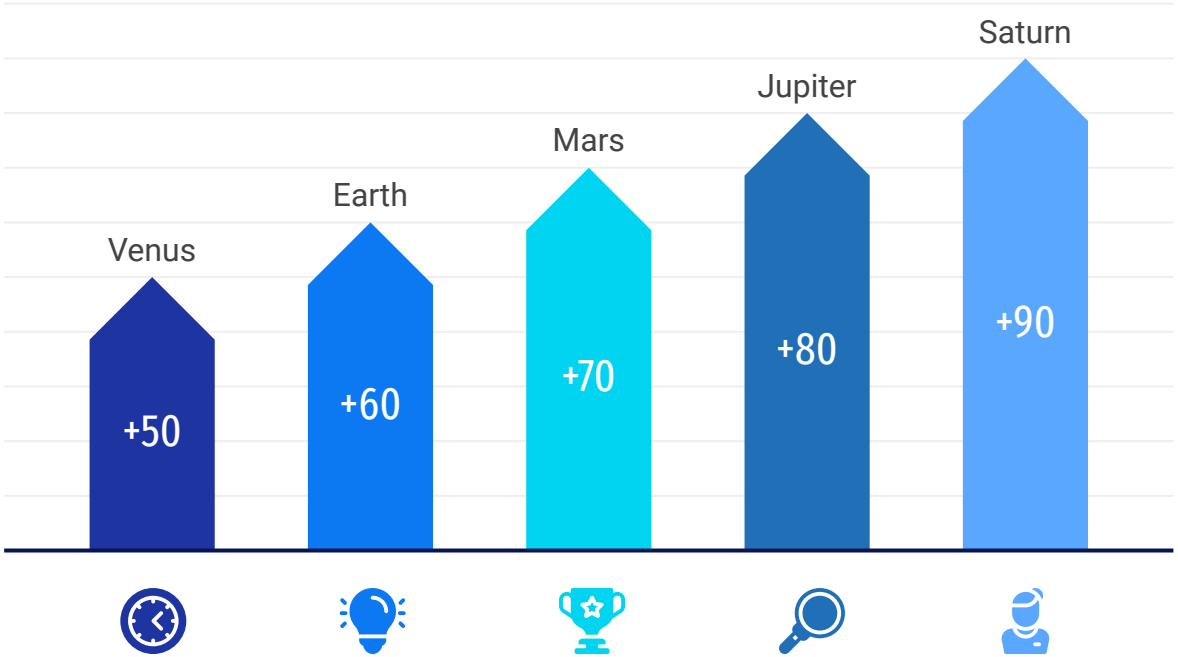
Saturn is composed of hydrogen and helium



Data Charts Infographics



Data Charts Infographics



Venus

Venus has a nice name and is the second planet from the Sun

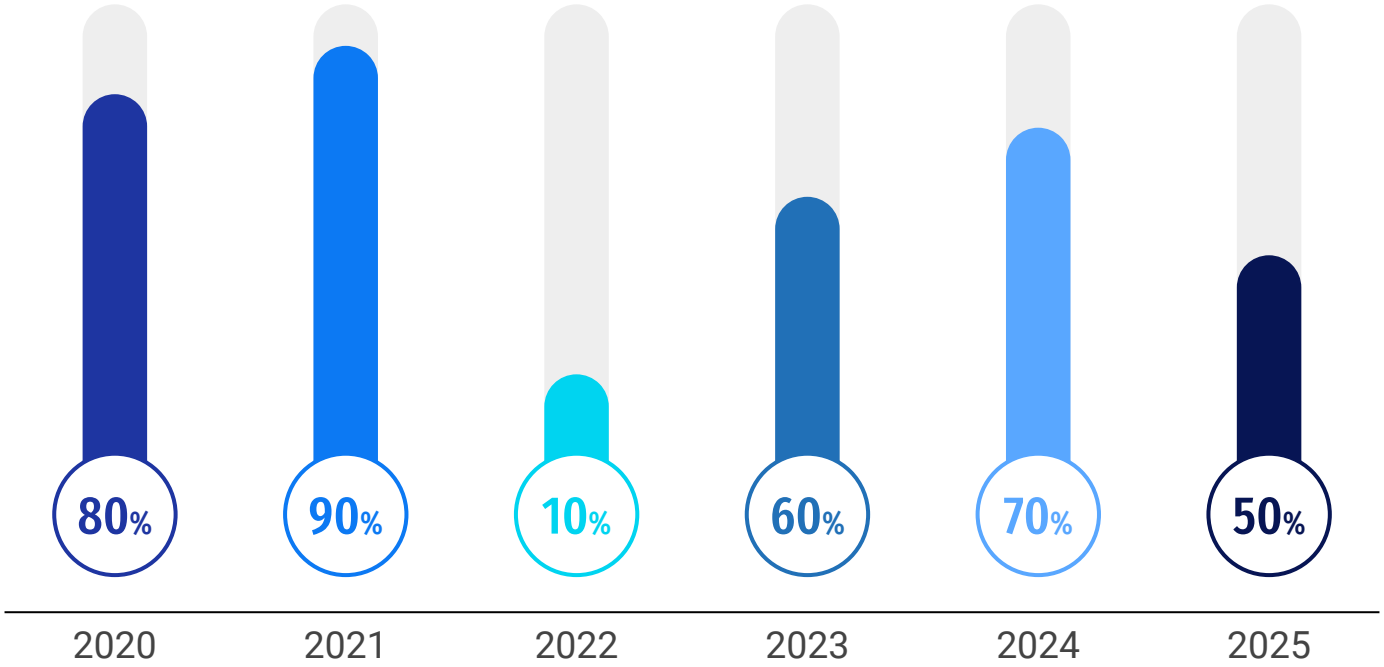
Saturn

Saturn is the ringed planet, composed of hydrogen and helium

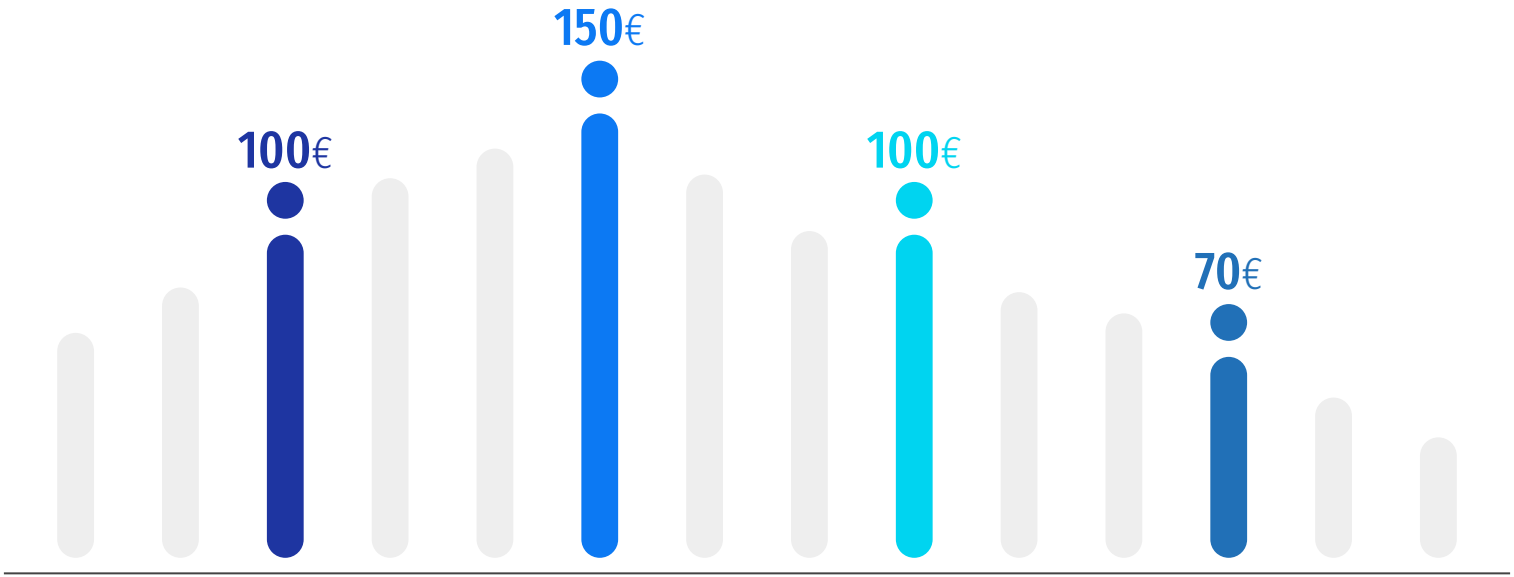
Mars

Despite being red, Mars is a cold planet full of iron oxide dust

Data Charts Infographics



Data Charts Infographics



Quarter 1

Venus has a nice name

Quarter 2

Mars is a cold place

Quarter 3

Jupiter is the biggest planet

Quarter 4

Saturn has several rings

Data Charts Infographics

Mercury

100%

Mercury is the smallest planet

Venus

50%

Venus has a beautiful name

Earth

60%

Earth in the only planet with life

Mars

30%

Mars is actually a cold place

Saturn

80%

Saturn is a gas giant and has rings

Data Charts Infographics

Mars

Despite being red, Mars is actually a cold place full of iron oxide dust

60%

Venus

Venus has a beautiful name, but it’s terribly hot, even hotter than Mercury

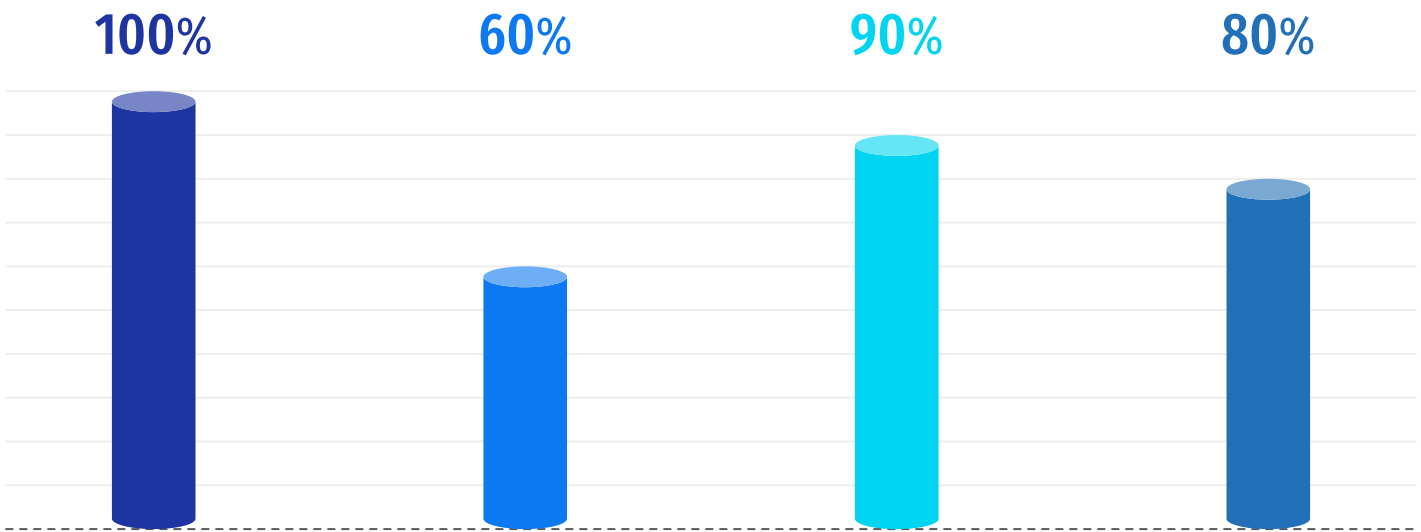
80%

Neptune

Neptune is the fourth-largest planet by diameter in the Solar System

50%

Data Charts Infographics



Venus

Venus has a beautiful name

Mars

Mars is actually a cold place

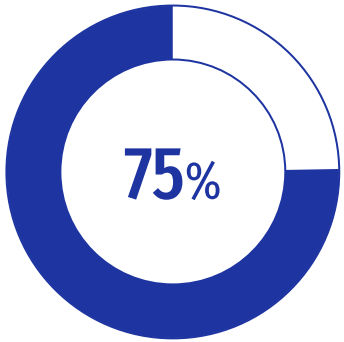
Mercury

Mercury is the smallest planet

Saturn

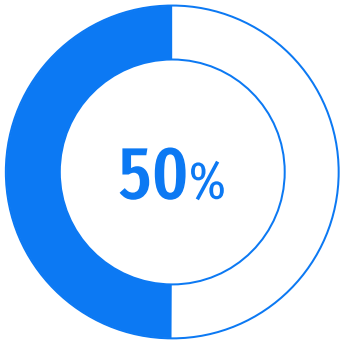
Saturn is a gas giant and has rings

Data Charts Infographics



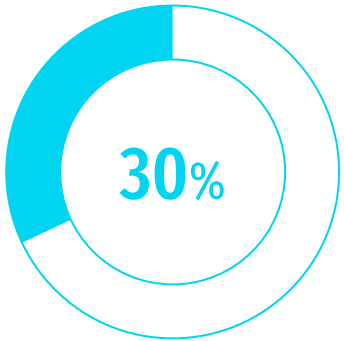
Saturn

It’s composed of hydrogen and helium



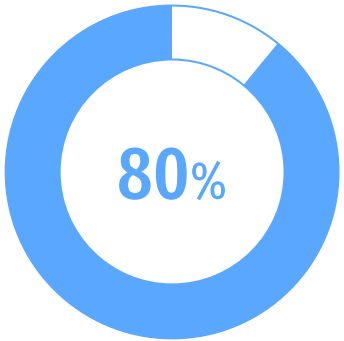
Mercury

Mercury is the closest planet to the Sun



Mars

Despite being red, Mars is a cold place



Jupiter

Jupiter is a gas giant and the biggest planet

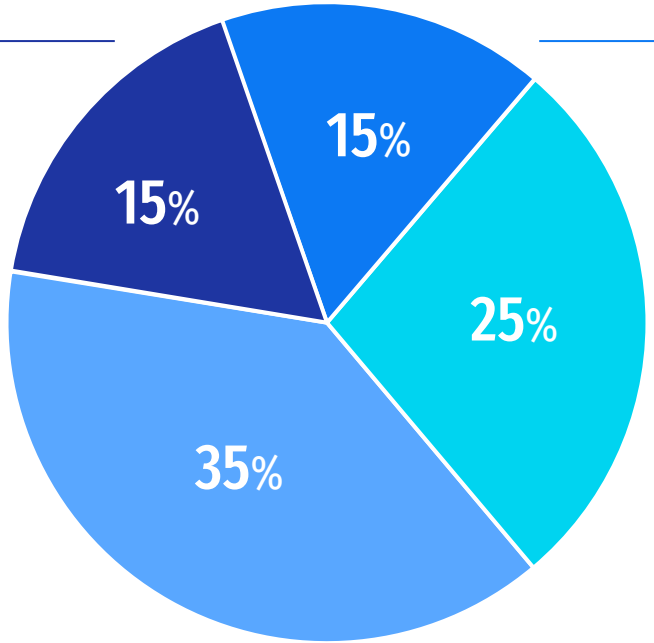
Data Charts Infographics

Quarter 1

Venus has a beautiful name and is the second planet from the Sun

Quarter 2

Despite being red, Mars is a cold place full of iron oxide dust



Quarter 3

Mercury is the closest planet to the Sun and the smallest one

Quarter 4

Jupiter is a gas giant and the biggest planet in the Solar System

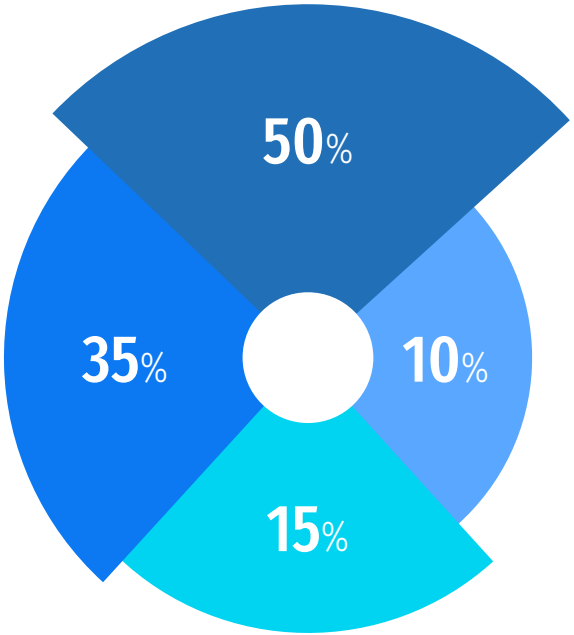
Data Charts Infographics

Venus

Venus has a beautiful name and is the second planet from the Sun

Mercury

Mercury is the closest planet to the Sun and the smallest one



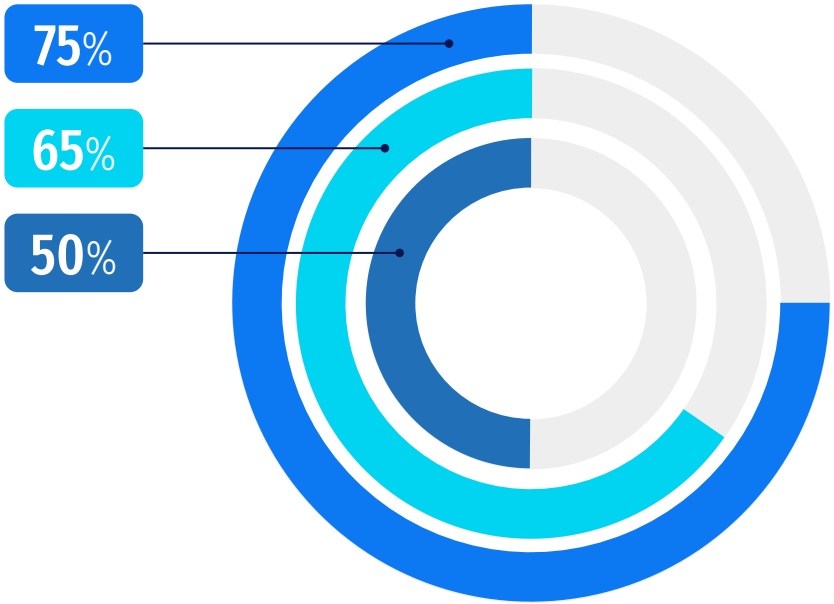
Mars

Despite being red, Mars is a cold place full of iron oxide dust

Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

Data Charts Infographics



Venus

Venus has a beautiful name and is the second planet from the Sun

Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

Mercury

Mercury is the closest planet to the Sun and the smallest of them all

Data Charts Infographics

50%

Venus

Venus has a beautiful name and is the second planet from the Sun

60%

Mars

Despite being red, Mars is a cold place full of iron oxide dust

70%

Mercury

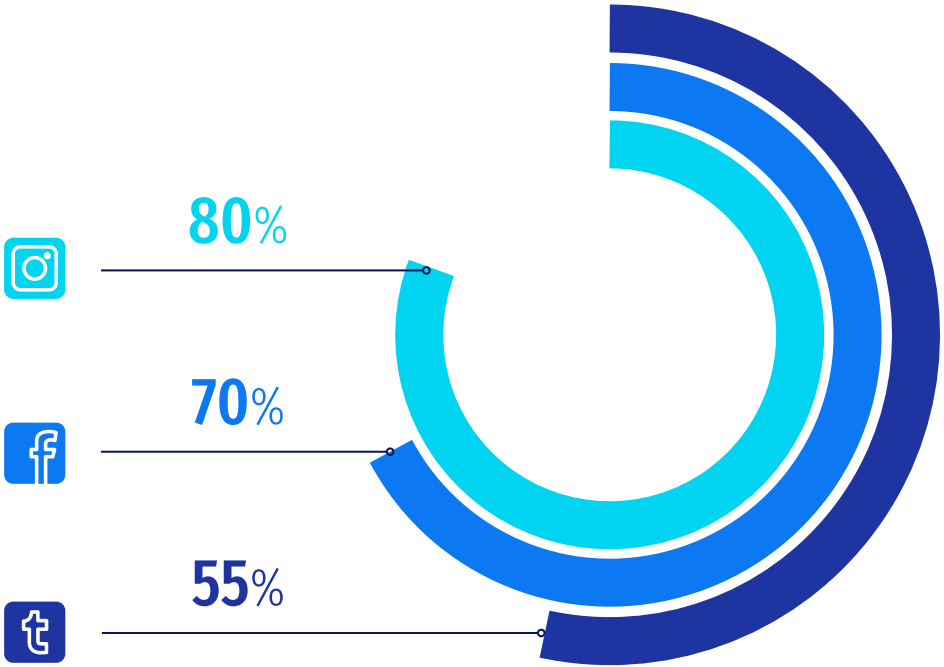
Mercury is the closest planet to the Sun and the smallest one

80%

Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

Data Charts Infographics



Venus

Venus has a beautiful name and is the second planet from the Sun

Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

Mercury

Mercury is the closest planet to the Sun and the smallest of them all

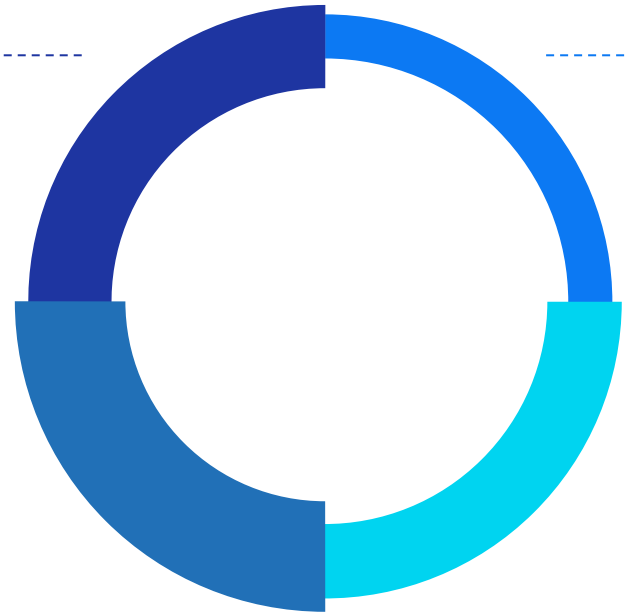
Data Charts Infographics

Venus

Venus has a beautiful name and is the second planet from the Sun

Mercury

Mercury is the closest planet to the Sun and the smallest one



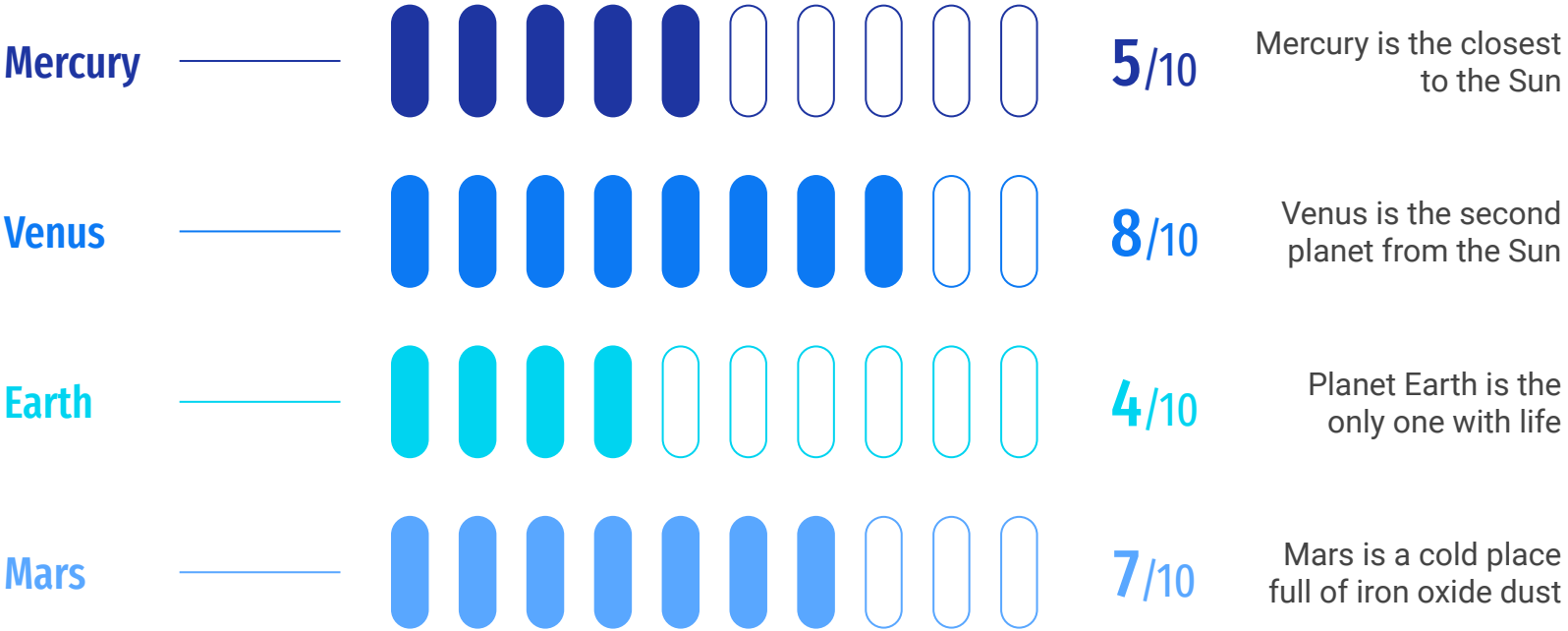
Mars

Despite being red, Mars is a cold place full of iron oxide dust

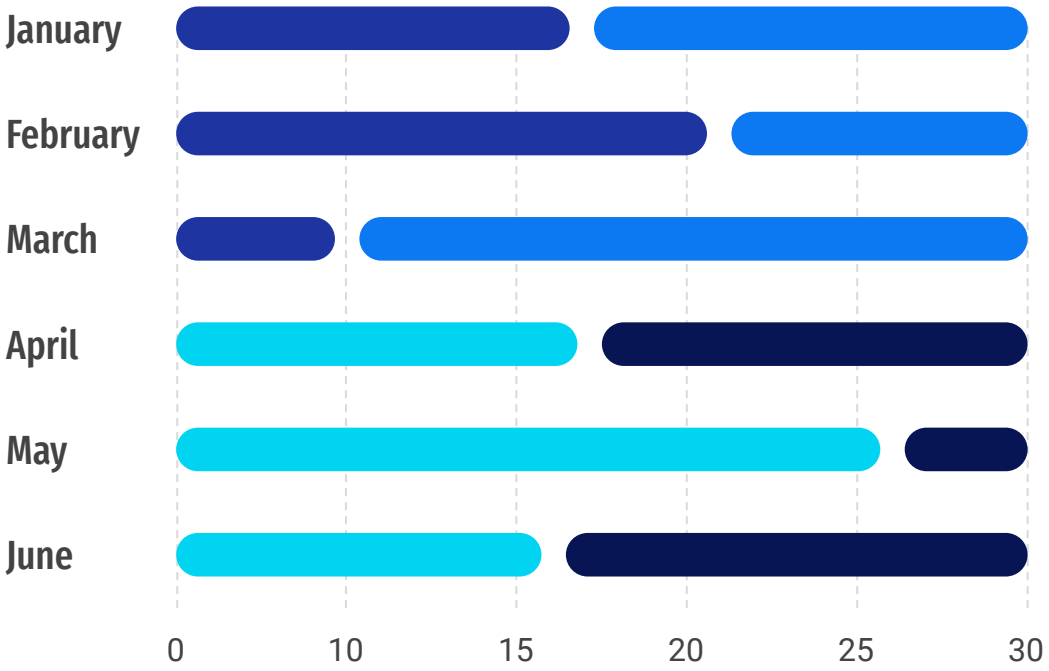
Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

Data Charts Infographics



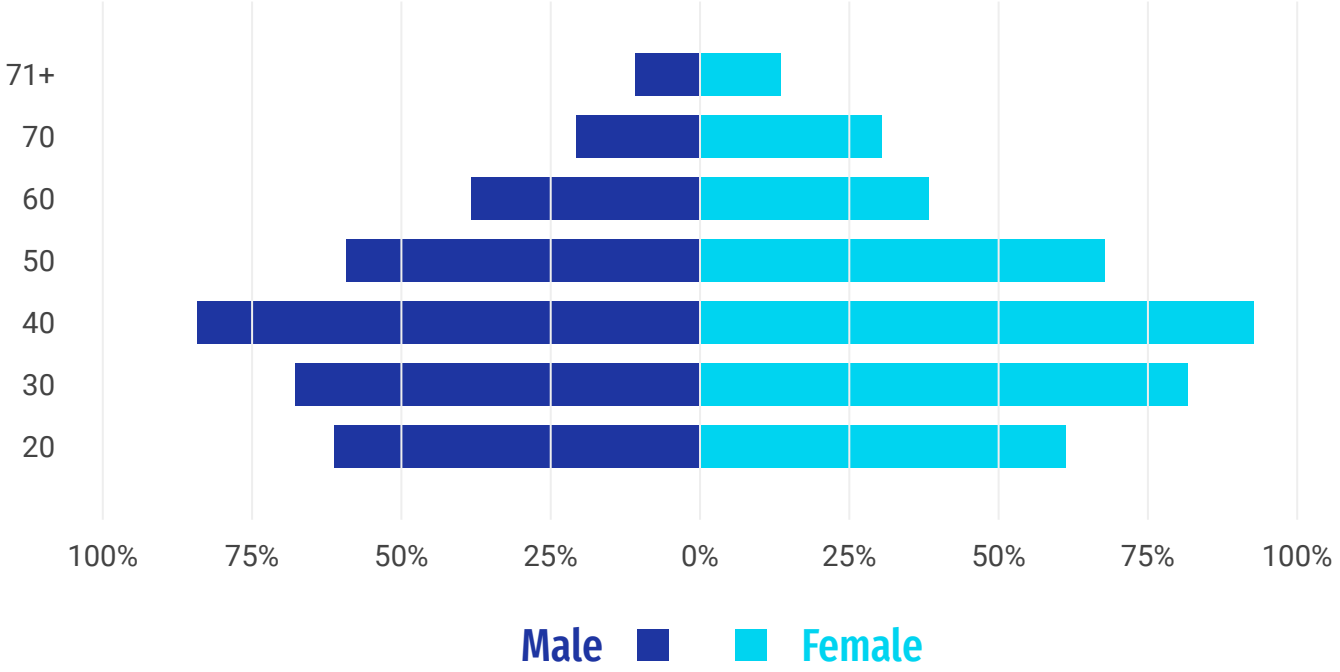
Data Charts Infographics



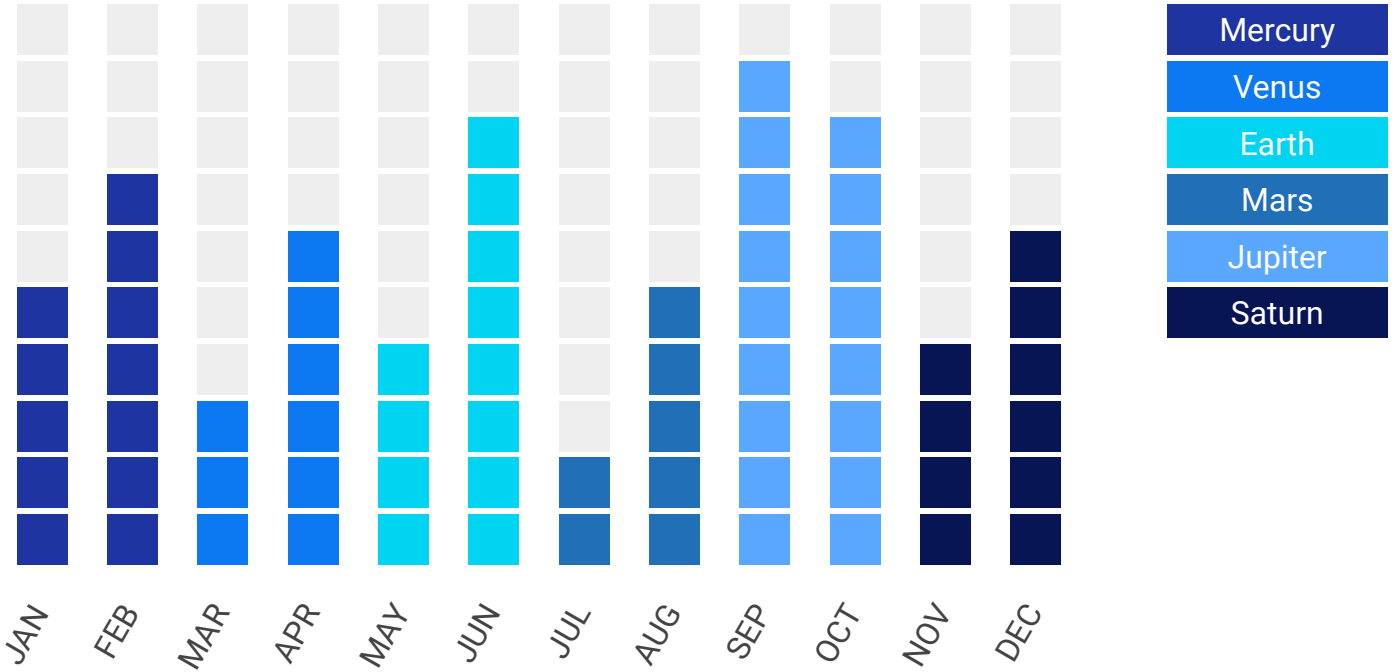
Planet Venus

Venus has a beautiful name and is the second planet from the Sun. It is terribly hot, even hotter than Mercury

Data Charts Infographics



Data Charts Infographics



Data Charts Infographics

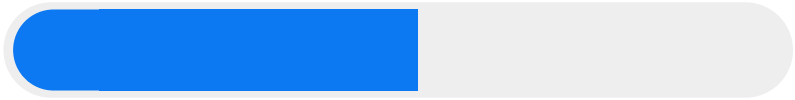
Venus

Venus has a beautiful name and is the second planet from the Sun

9,000

Mercury

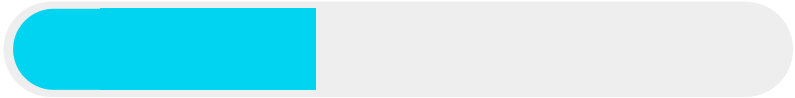
Mercury is the closest planet to the Sun and the smallest one



5,000

Mars

Despite being red, Mars is a cold place full of iron oxide dust



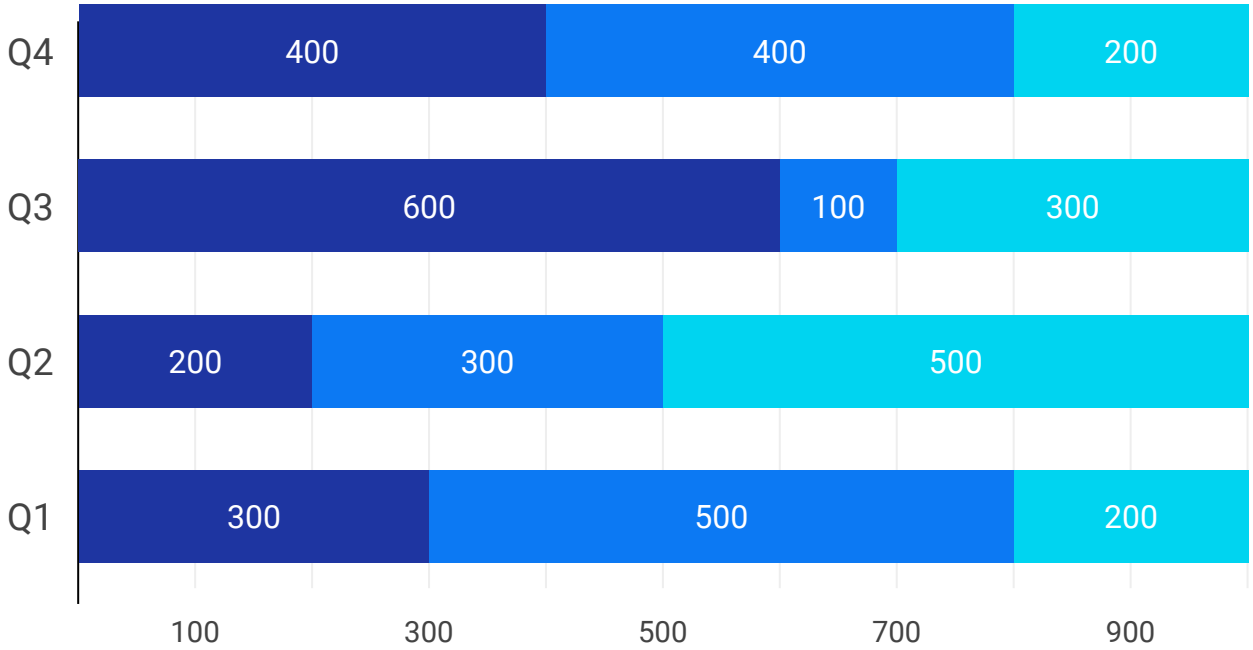
4,000

Jupiter

Jupiter is a gas giant and the biggest planet in the Solar System

8,000

Data Charts Infographics



Mercury

Mercury is the smallest planet

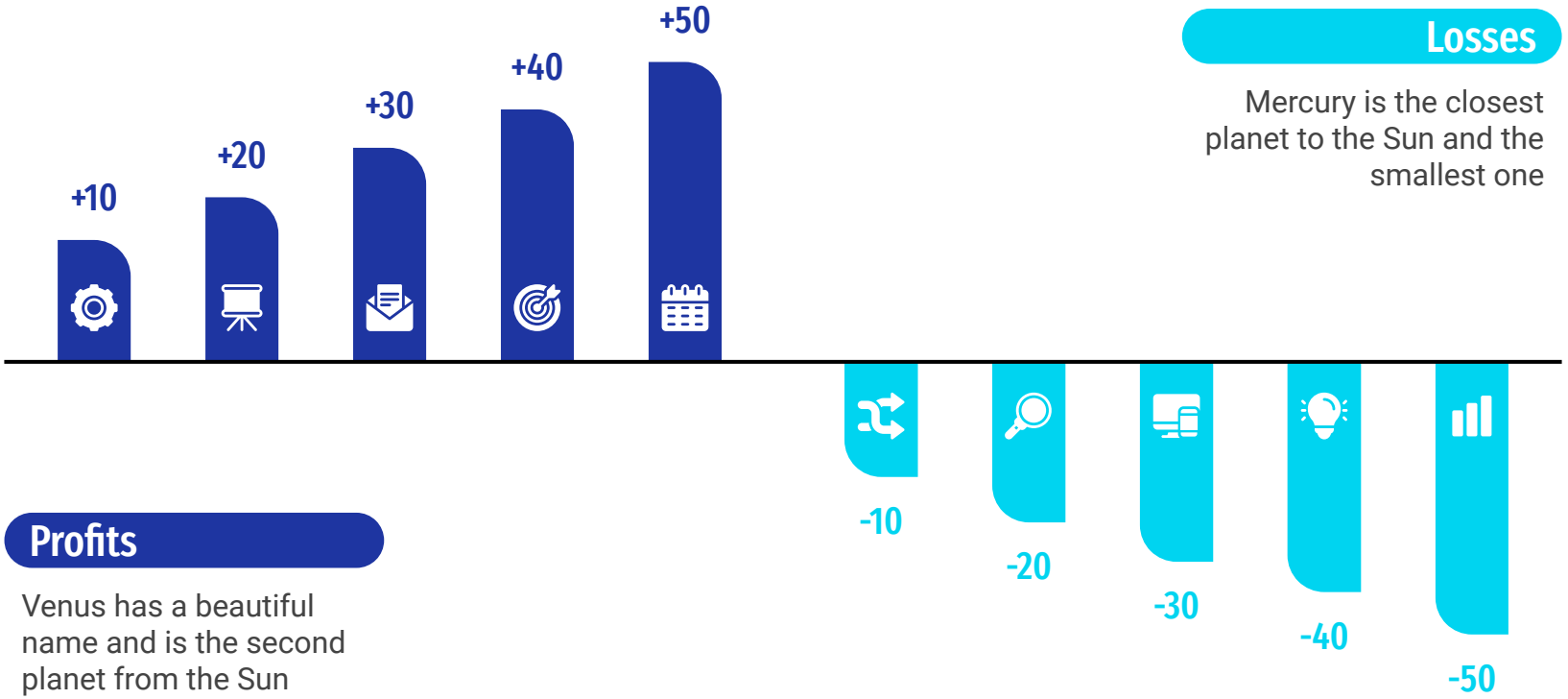
Venus

Venus is the second planet from the Sun

Earth

Earth is the only planet with life

Data Charts Infographics



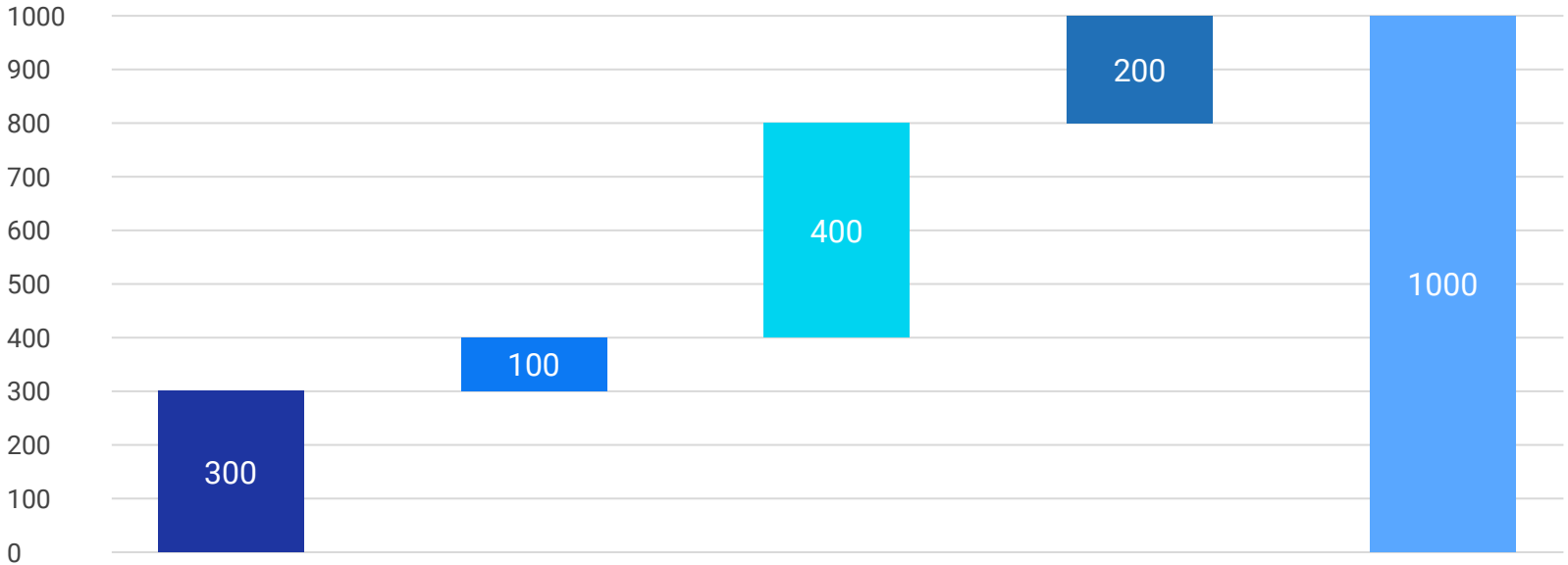
Losses

Mercury is the closest planet to the Sun and the smallest one

Profits

Venus has a beautiful name and is the second planet from the Sun

Data Charts Infographics



Venus

Venus has a beautiful name

Mars

Mars is a cold place

Jupiter

Jupiter is the biggest planet

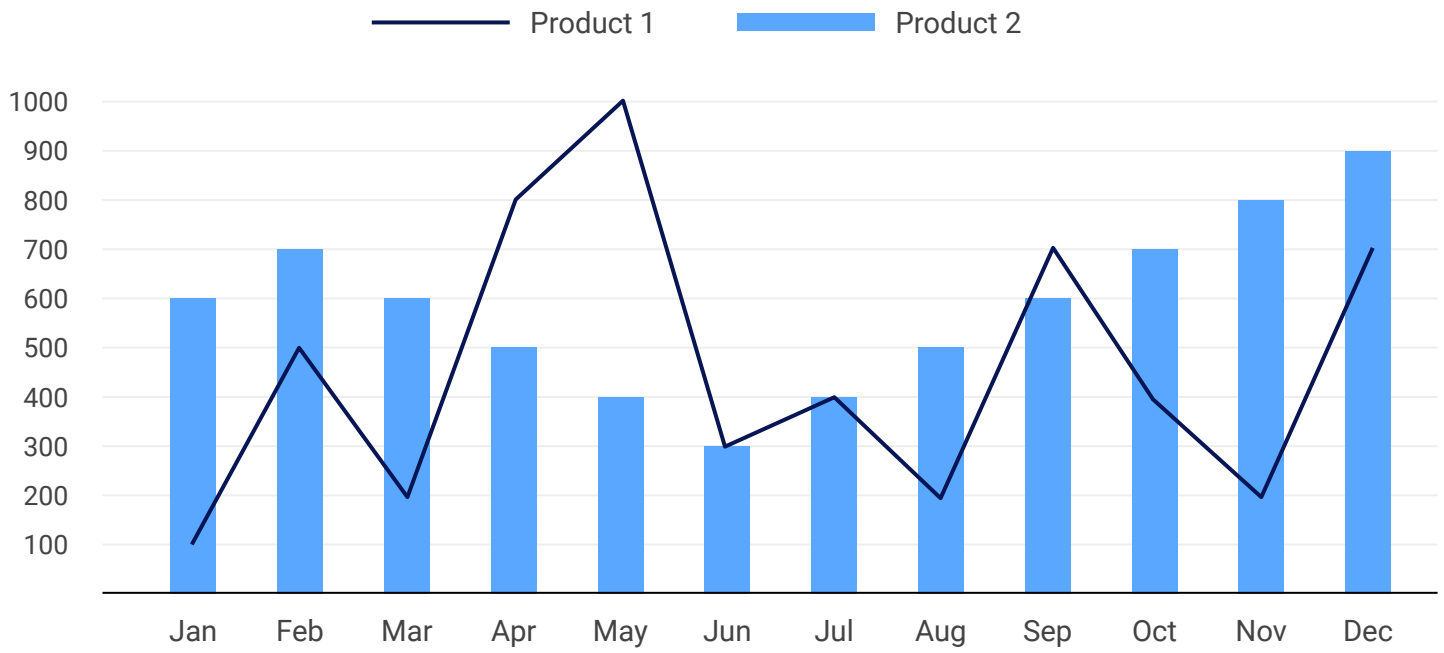
Saturn

Saturn has several rings

Earth

Earth is the planet with life

Data Charts Infographics



Data Charts Infographics

Result 1

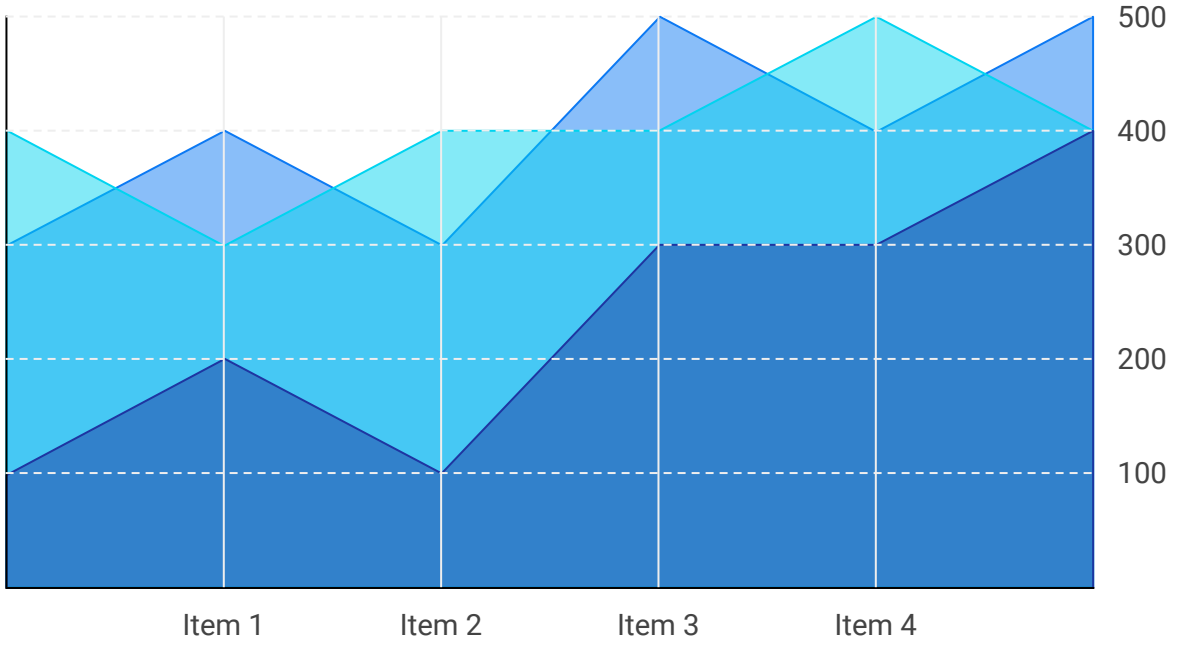
Venus is the second planet from the Sun

Result 2

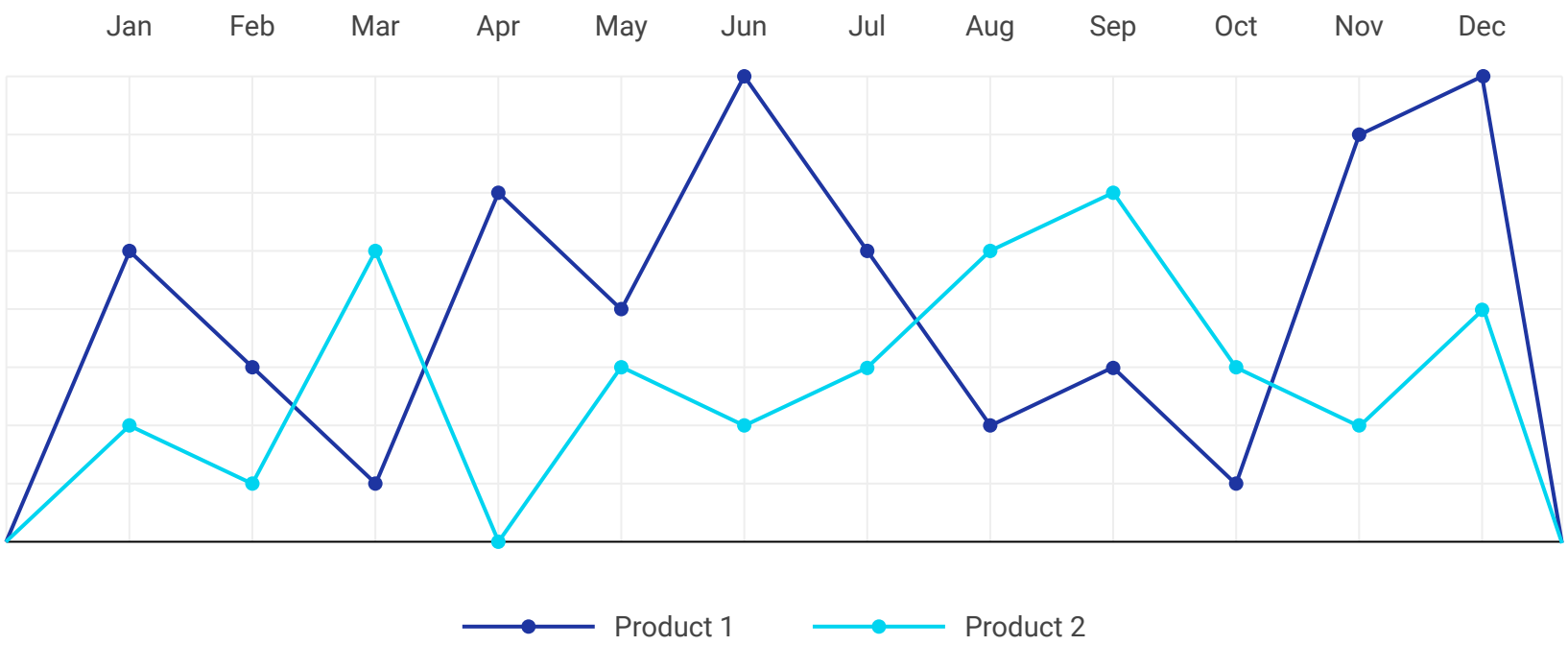
Saturn is composed of hydrogen and helium

Result 3

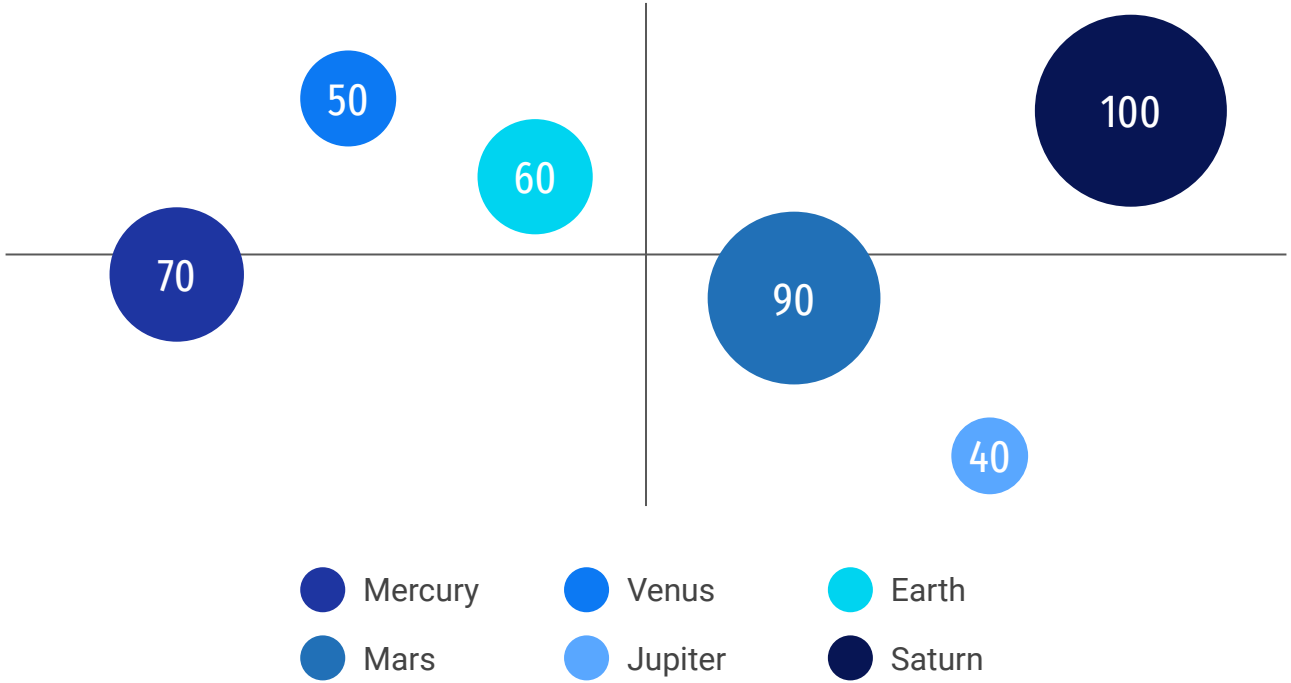
Mars is a cold planet full of iron oxide dust



Data Charts Infographics



Data Charts Infographics



Instructions for use (free users)

In order to use this template, you must credit **Slidesgo** by keeping the Thanks slide.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial purposes.

You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute this Slidesgo Template (or a modified version of this Slidesgo Template) or include it in a database or in any other product or service that offers downloadable images, icons or presentations that may be subject to distribution or resale.
- Use any of the elements that are part of this Slidesgo Template in an isolated and separated way from this Template.
- Delete the “Thanks” or “Credits” slide.
- Register any of the elements that are part of this template as a trademark or logo, or register it as a work in an intellectual property registry or similar.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

https://slidesgo.com/faqs and https://slidesgo.com/slidesgo-school

Instructions for use (premium users)

In order to use this template, you must be a Premium user on **Slidesgo**.

You are allowed to:

- Modify this template.
- Use it for both personal and commercial purposes.
- Hide or delete the “Thanks” slide and the mention to Slidesgo in the credits.
- Share this template in an editable format with people who are not part of your team.

You are not allowed to:

- Sublicense, sell or rent this Slidesgo Template (or a modified version of this Slidesgo Template).
- Distribute this Slidesgo Template (or a modified version of this Slidesgo Template) or include it in a database or in any other product or service that offers downloadable images, icons or presentations that may be subject to distribution or resale.
- Use any of the elements that are part of this Slidesgo Template in an isolated and separated way from this Template.
- Register any of the elements that are part of this template as a trademark or logo, or register it as a work in an intellectual property registry or similar.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

https://slidesgo.com/faqs and https://slidesgo.com/slidesgo-school

Infographics

You can add and edit some **infographics** to your presentation to present your data in a visual way.

- Choose your favourite infographic and insert it in your presentation using Ctrl C + Ctrl V or Cmd C + Cmd V in Mac.
- Select one of the parts and **ungroup** it by right-clicking and choosing “Ungroup”.
- **Change the color** by clicking on the paint bucket.
- Then **resize** the element by clicking and dragging one of the square-shaped points of its bounding box (the cursor should look like a double-headed arrow). Remember to hold Shift while dragging to keep the proportions.
- **Group** the elements again by selecting them, right-clicking and choosing “Group”.
- Repeat the steps above with the other parts and when you’re done editing, copy the end result and paste it into your presentation.
- Remember to choose the “**Keep source formatting**” option so that it keeps the design. For more info, please visit **Slidesgo School**.