

Library Academy

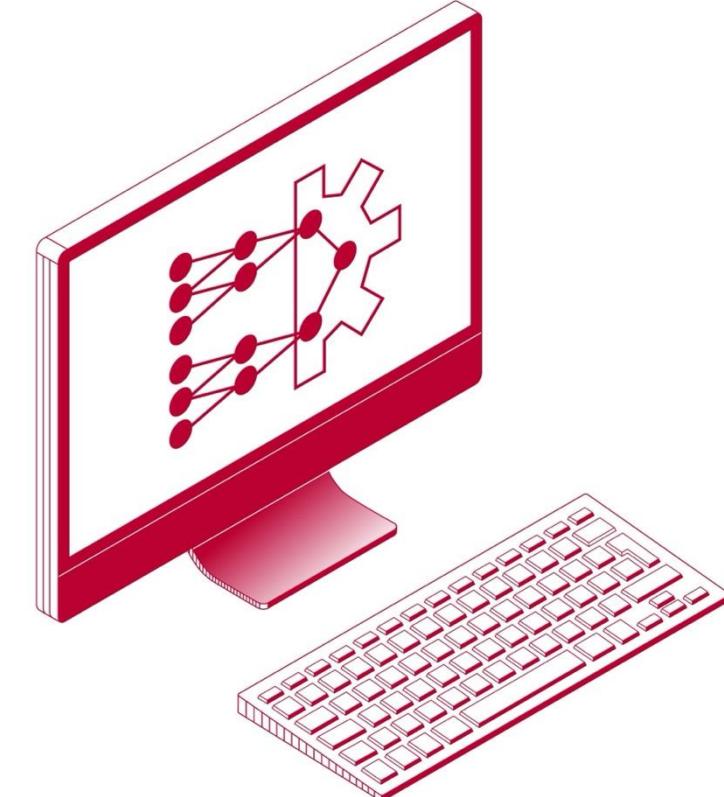
Innovate your education

21-01-2026



Ontdek wat data vertellen: Aan de slag met Machine Learning in JASP

Don van den Bergh & Johnny van Doorn
Library Academy 2026



Outline

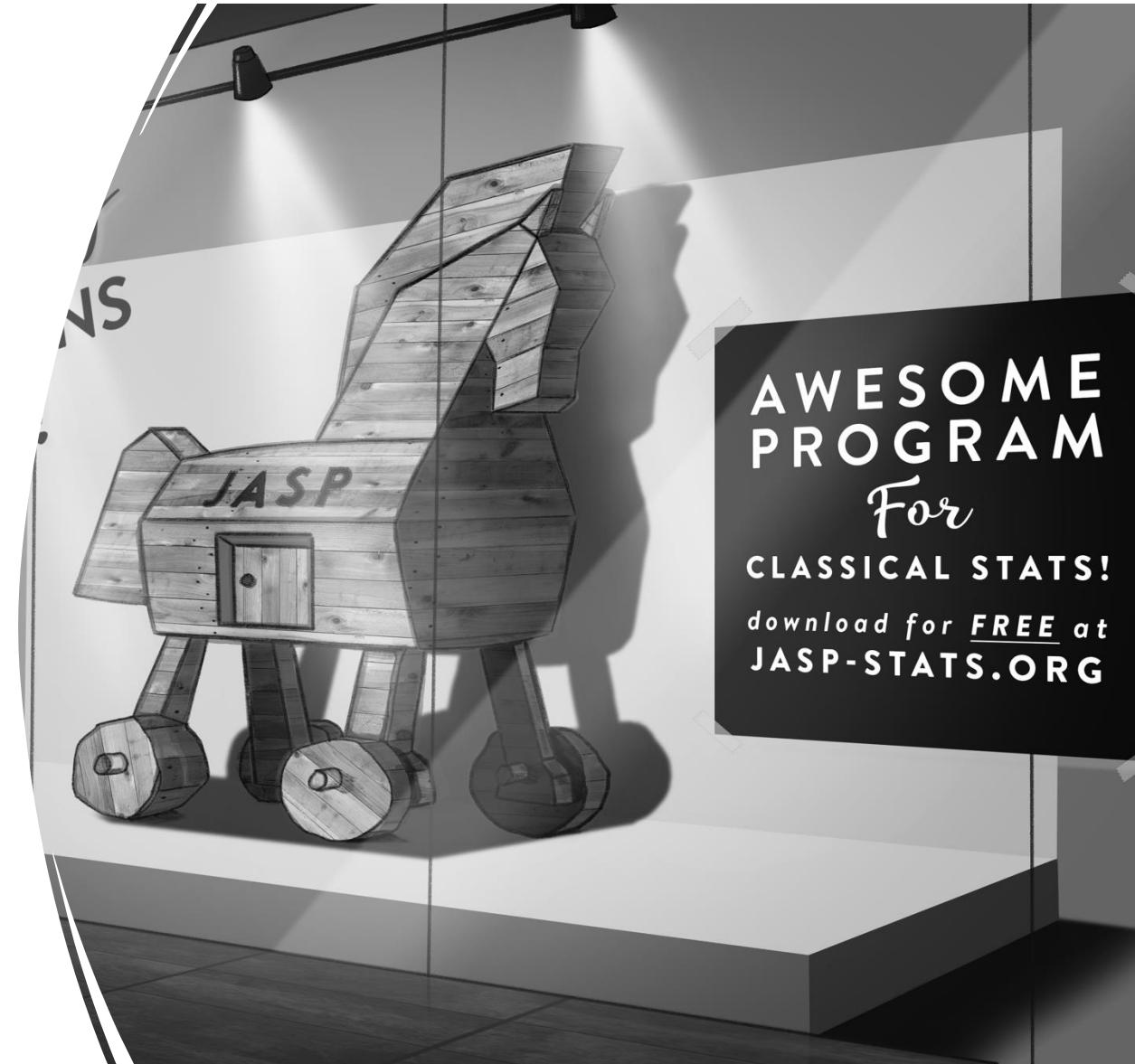
- Basics
 - What is JASP?
 - Regression
- Machine Learning
 - General philosophy
 - K-nearest neighbors
- Exercises



edu.nl/aa8ch

What is JASP?

- Developed at UvA over the past 10 years, funded by research grants (NWO/EU)
- Graphical user interface for conducting frequentist and Bayesian statistics
- <https://jasp-stats.org/>
- Open-source → forever free!



What is JASP?

The screenshot shows the JASP software interface. The top navigation bar includes icons for Edit Data, Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, and Factor. The main window is focused on the ANOVA module. On the left, the 'ANOVA' settings panel shows 'Dependent Variable' set to 'Score' and 'Fixed Factors' set to 'Group'. Below these are sections for 'WLS Weights' and 'Display' (with 'Estimates of effect size' checked). On the right, the 'ANOVA' results panel displays a descriptive text about the dataset, followed by an ANOVA table and a Descriptives plot.

ANOVA

Dependent Variable: Score

Fixed Factors: Group

WLS Weights:

Display:

- Descriptive statistics
- Estimates of effect size
 - ω^2
 - partial ω^2
 - η^2
 - partial η^2
 - Confidence intervals 95 %
- Vovk-Sellke maximum p-ratio

Model

Assumption Checks

Contrasts

Order Restricted Hypotheses

Post Hoc Tests

ANOVA

Description:
This data set, "Response to Eye Color", provides post-advertisement attitudes towards a brand expressed by four different groups - each group saw the same advertisement except for the aspect that was manipulated: the eye-color of the model.

Variables:

- **Group** - Experimental conditions ('Blue' = Model with blue eyes, 'Brown' = Model with brown eyes, 'Green'= Model with green eyes, 'Down' = Model's eye color cannot be seen).
- **Subj** - Participant number.
- **Score** - An average of 10 survey questions about attitudes towards the brand (7-point Likert scale). Higher averages correspond to more positive attitudes.

This example JASP file demonstrates the use of one-way ANOVA. Specifically, we assess the adequacy of the null hypothesis that the attitudes are the same regardless of the eye-color of the model

ANOVA - Score

	Cases	Sum of Squares	df	Mean Square	F	p	ω^2	95% CI for ω^2	
Group		24.420	3	8.140	2.894	.036	0.025	0.000	0.069
Residuals		613.139	218	2.813					

Note. Type III Sum of Squares

Descriptives

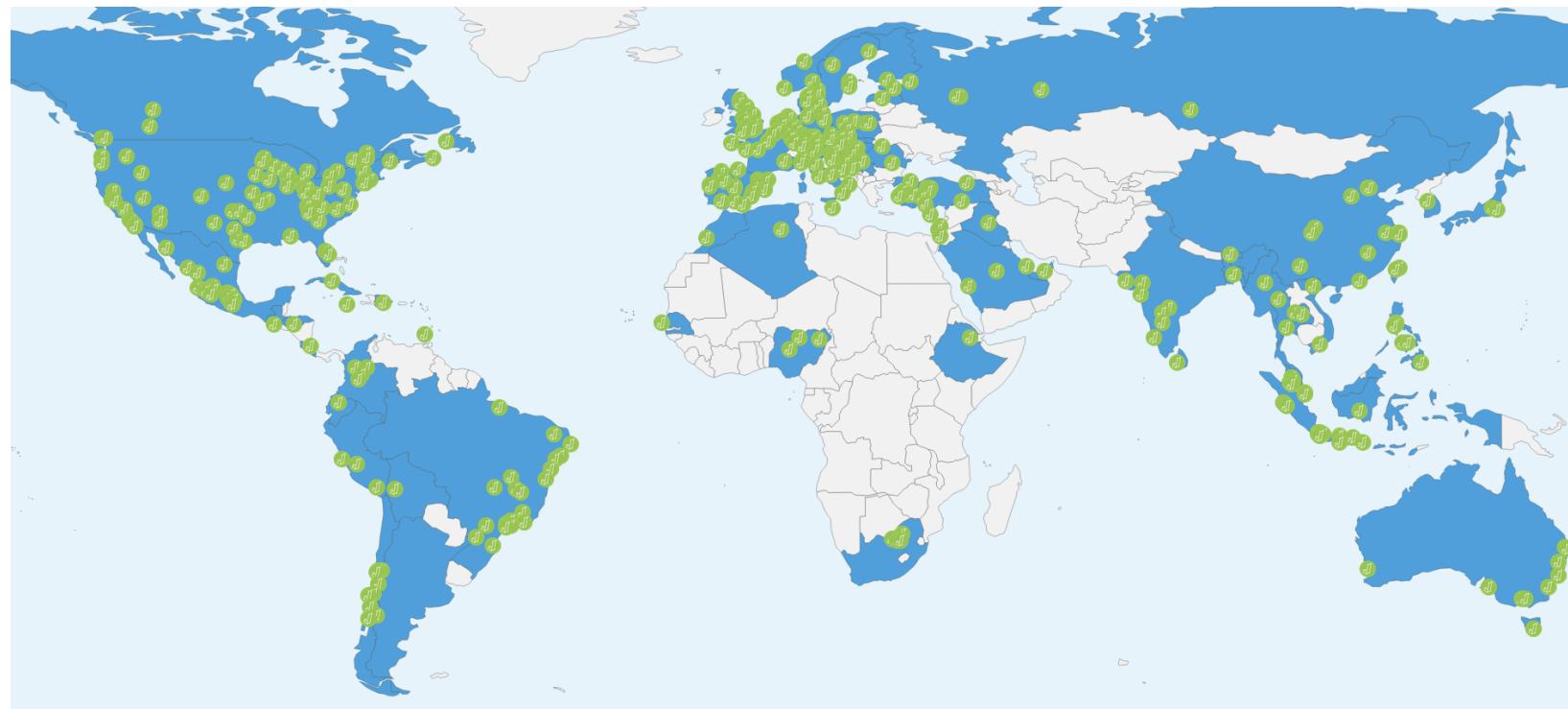
Descriptives plots

A line plot titled 'Descriptives plots' showing the mean score for each group. The x-axis categories are 'Down', 'Blue', 'Brown', and 'Green'. The y-axis ranges from 2.5 to 4.5. Error bars are shown for each point.

Group	Mean Score	Lower CI	Upper CI
Down	3.1	2.6	3.6
Blue	3.2	2.8	3.6
Brown	3.7	3.2	4.2
Green	3.8	3.4	4.2

What is JASP?

- Used at 374 universities across 76 countries
- 100,000 monthly downloads



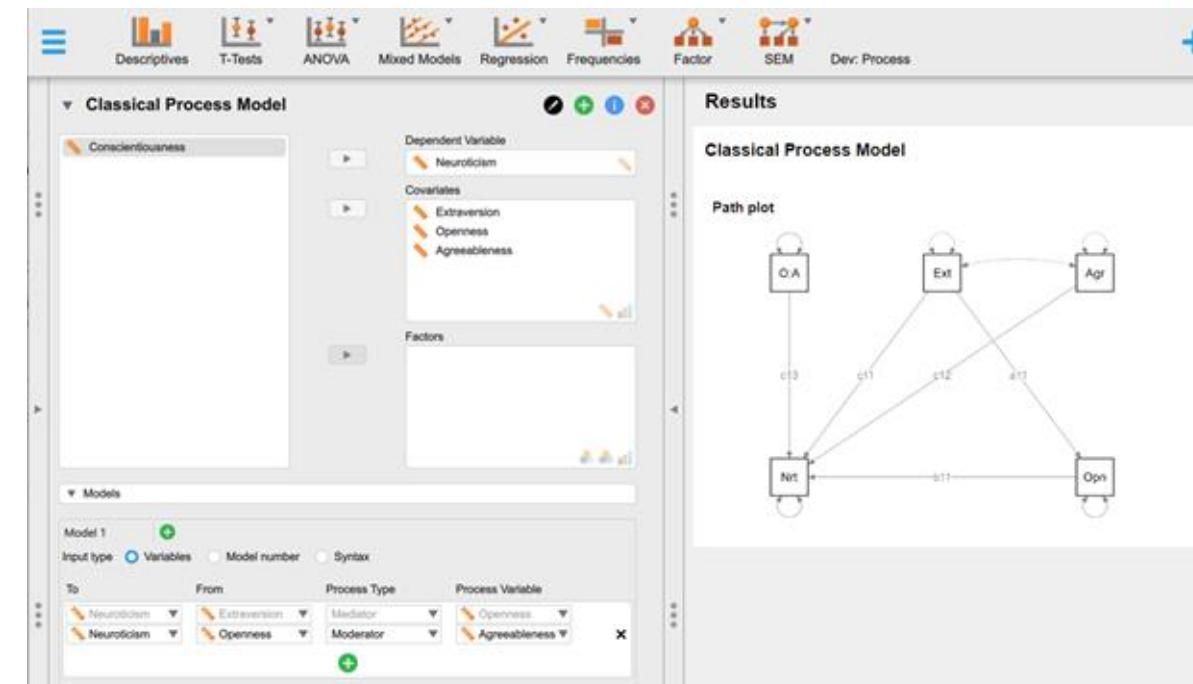
<https://jasp-stats.org/teaching-with-jasp/#worldmap>

Features

- [Website overview](#)
- [JASP vs. SPSS feature comparison](#)
- Data formats: .sav, .xls, .txt, .csv, .ods, .tsv, .dta, .por, .sas7bdat, .sas7bcat, and the .jasp format
- APA tables
- OSF integration
- R console
- Compute columns
- Filtering

Feature Roadmap

- Full syntax mode ([blog about the first implementation](#))
- Better data manipulation
- Select filters



Other Handy Resources

- [The JASP Video Library](#)
- [How to Use JASP – Inventory of blogs/videos/gifs for frequentist and Bayesian analyses](#)
- [JASP YouTube page](#)
- [Step By Step Guide: 1. Bayesian One-Way ANOVA](#) and the [full playlist](#)
- JASP on Bluesky - <https://bsky.app/profile/jaspstats.bsky.social>
- JASP forum - <https://forum.cogsci.nl/index.php?p=/categories/jasp-bayesfactor>
- Found a bug? Please report on Github: <https://github.com/jasp-stats/jasp-issues/issues>
- [JASP Verification Project](#)
- More JASP workshops: <https://jasp-stats.org/workshop/>

JASP Literature

- [The JASP Data Library](#)
- [Discovering Statistics Using JASP](#)
- [Learning Statistics with JASP: A Tutorial for Psychology Students and Other Beginners by Danielle J. Navarro, David R. Foxcroft, and Thomas J. Faulkenberry](#)
- [Statistics of Doom by Erin Buchanan](#)
- [Statistical Analysis in JASP. A Guide for Students by Mark Goss-Sampson](#)
- [Quantitative Analysis with JASP open-source software by Chris Halter](#) (amazon)



| JASP

Data Management

The screenshot shows a data management interface with a toolbar at the top and a table below. The toolbar includes icons for Analyses, Synchronisation, Resize Data, Insert, and Remove. The table displays five rows of data about metal band members, with columns for Name, Instrument, Current member, and Headbanging intensity.

	Name	Instrument	Current member	Headbanging intensity
1	Lars Ulrich	Drums	Yes	1 Light
2	James Hetfield	Guitar	Yes	1 Heavy
3	Kirk Hammett	Guitar	Yes	1 Light
4	Rob Trujillo	Bass	Yes	1 Moderate
5	Jason Newsted	Bass	No	0 Heavy

The Variable View

Name: Name Long name: Full name of Metallica band member

Column type: Nominal Description: ...

Computed type: Not computed

Label editor Missing values

Filter	Value	Label
✓	Lars Ulrich	Lars Ulrich
✓	James Hetfield	James Hetfield
✓	Kirk Hammett	Kirk Hammett
✓	Rob Trujillo	Rob Trujillo
✓	Jason Newsted	Jason Newsted

1 N
1 N
↑↓
▲▲
▼▼

Variable Types

- Scale 
 - Numbers (e.g., 7, 0, 120, 8.5)
- Nominal 
 - Categories (e.g., ‘Control group’, ‘Experimental group’)
- Ordinal 
 - Ordered values (e.g., ‘Dislike’, ‘Neutral’, ‘Like’)

Variable Settings

Name: Headbanging intensity

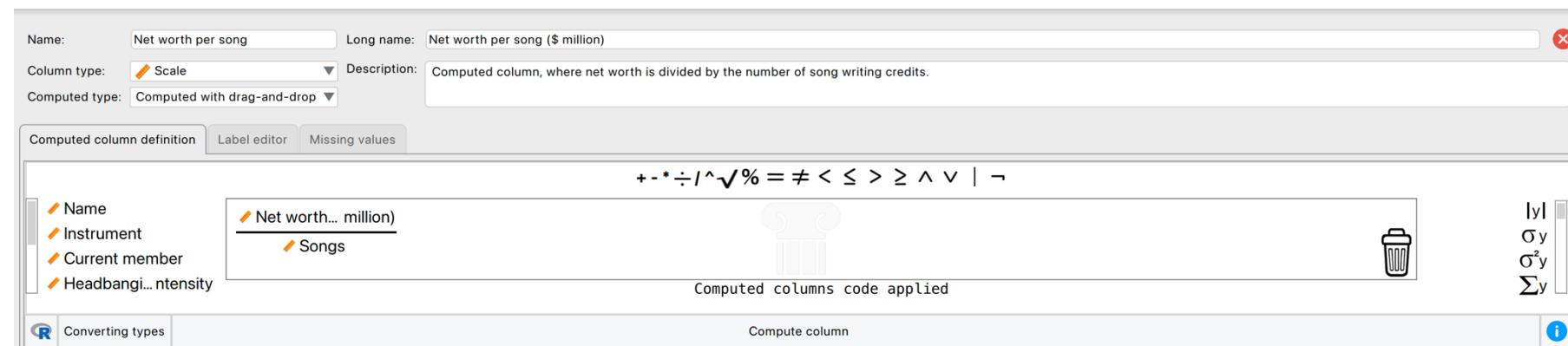
Column type: Ordinal

Computed type: Not computed

Label editor Missing values

Filter	Value	Label
✓	1	Light
✓	2	Moderate
✓	3	Heavy

Computing a New Variable



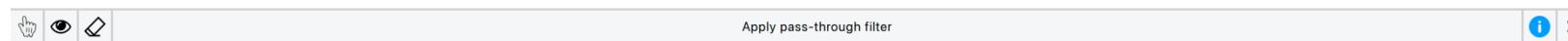
Filtering Data

- Using Variable Settings
- Using the Filter functionality
 - Drag and drop

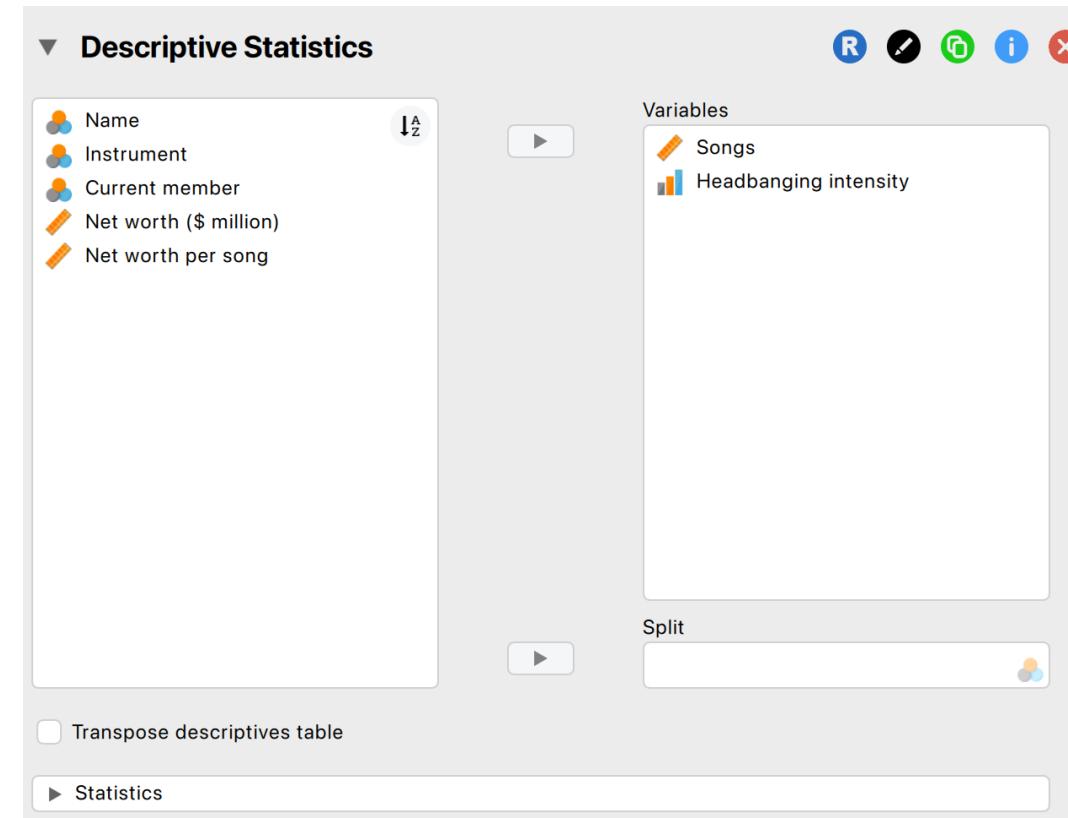


- R-mode

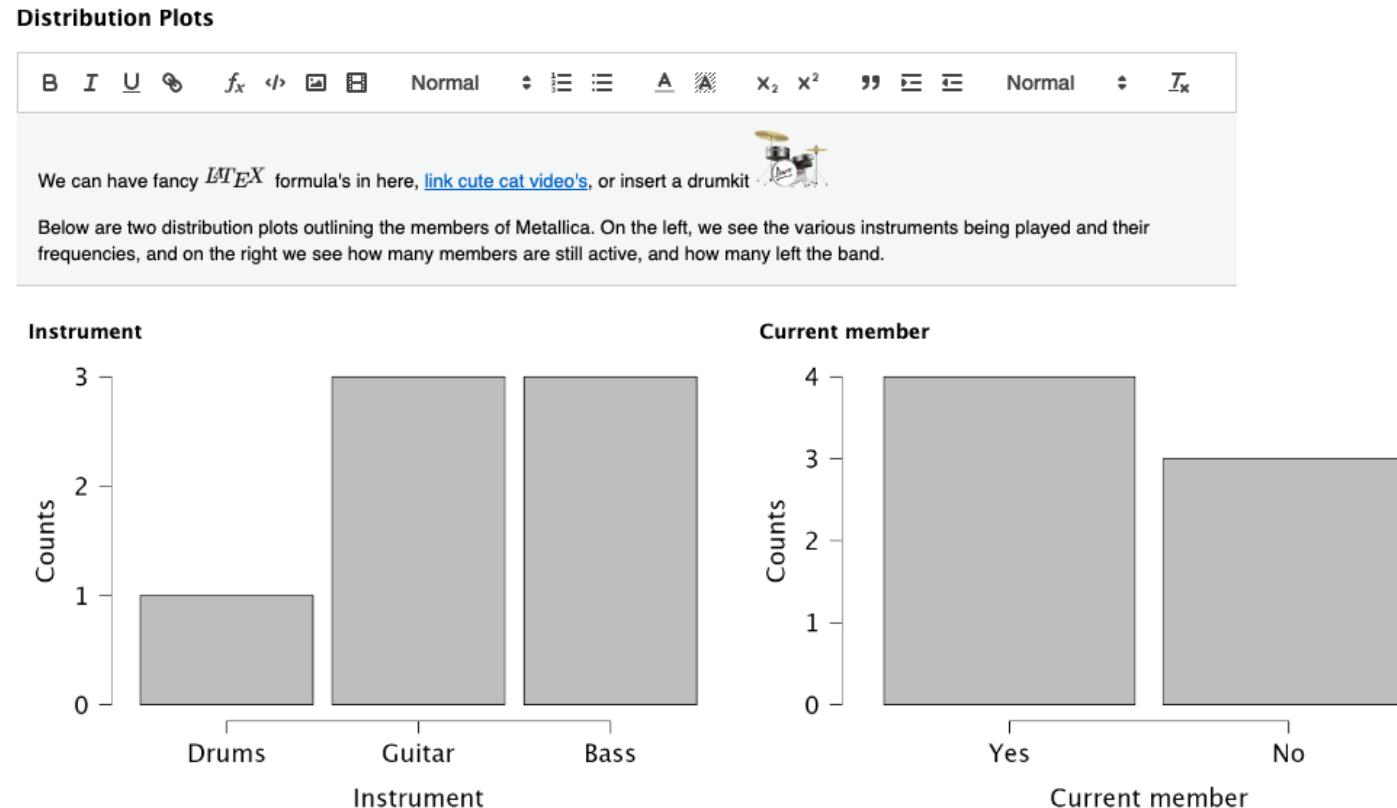
Songs > 100



Descriptives – input window



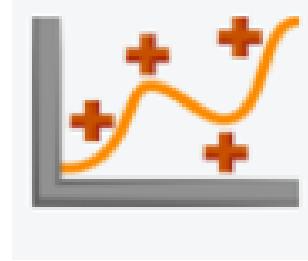
Descriptives – output window



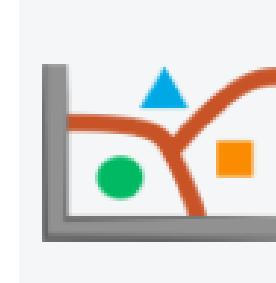


| JASP

There are lots of different machine learning algorithms

**Regression**

- Boosting
- Decision tree
- K-nearest neighbors
- Linear
- Neural network
- Random forest
- Regularized linear
- Support vector machine

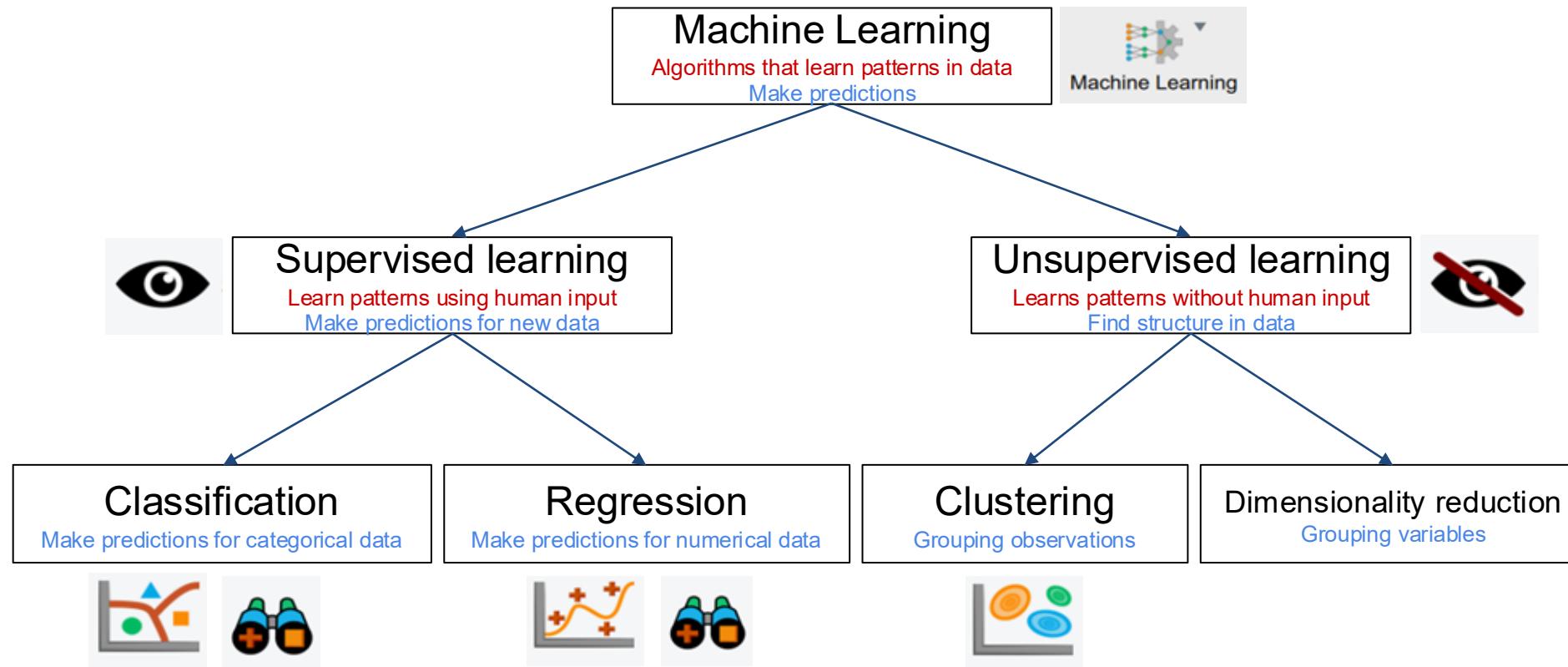
**Classification**

- Boosting
- Decision tree
- K-nearest neighbors
- Linear discriminant
- Logistic/Multinomial
- Naive Bayes
- Neural network
- Random forest
- Support vector machine

**Clustering**

- Density-based
- Fuzzy c-means
- Hierarchical
- Model-based
- Neighborhood-based
- Random forest

Machine learning is about finding patterns, the goal is making predictions



- Some algorithms are ‘black boxes’



- Understanding all algorithms is not the most important, understanding how to evaluate their results is

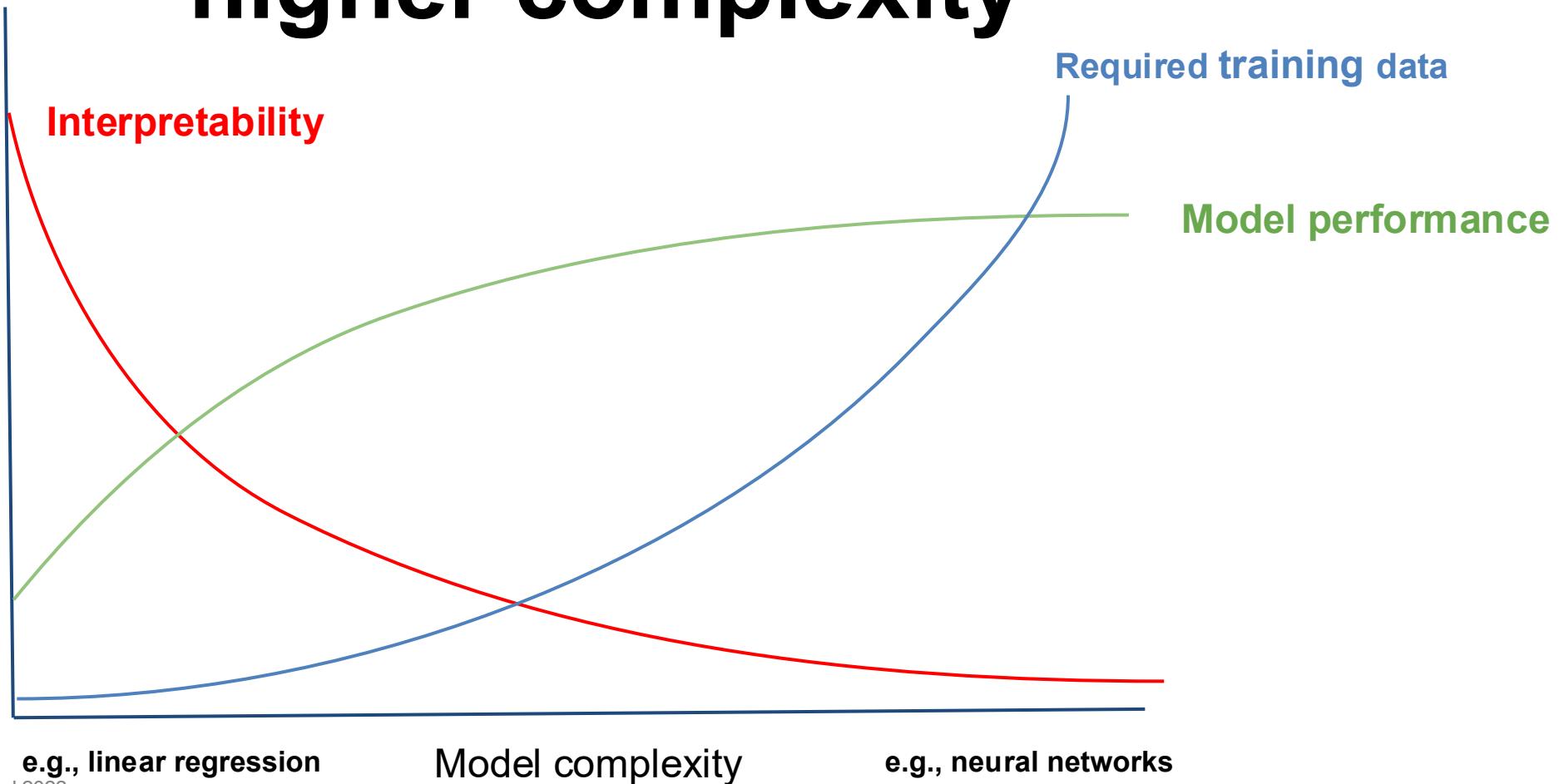
Explainable

Decision tree Linear regression
K-nearest neighbors

Not explainable

Support vector machines
Regularized linear regression
Random forests
Boosting
Neural network

Better performance (often) comes with higher complexity



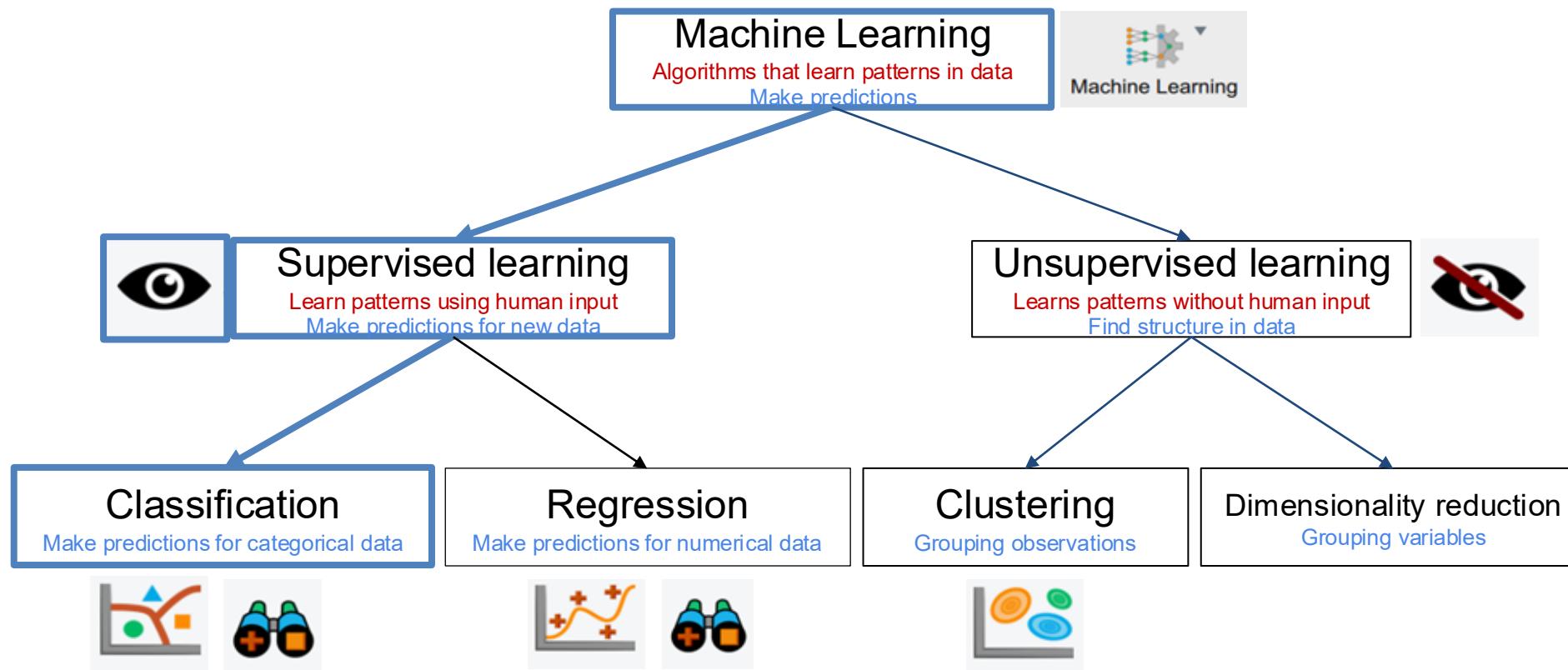
Learning objectives for today

- Train machine learning algorithms (regression, classification, clustering) using JASP
- Understand how to evaluate the quality of these algorithms
- Apply these algorithms to predict new data

Learning objectives for today

- Train machine learning algorithms (regression, **classification**, clustering) using JASP
- Understand how to evaluate the quality of these algorithms
- Apply these algorithms to predict new data

Machine learning is about finding patterns, the goal is making predictions



A typical classification classification workflow

1. Training the classification model
2. Optimizing the model (optional)
3. Evaluating the quality of the predictions of the model
4. Apply the model to new data

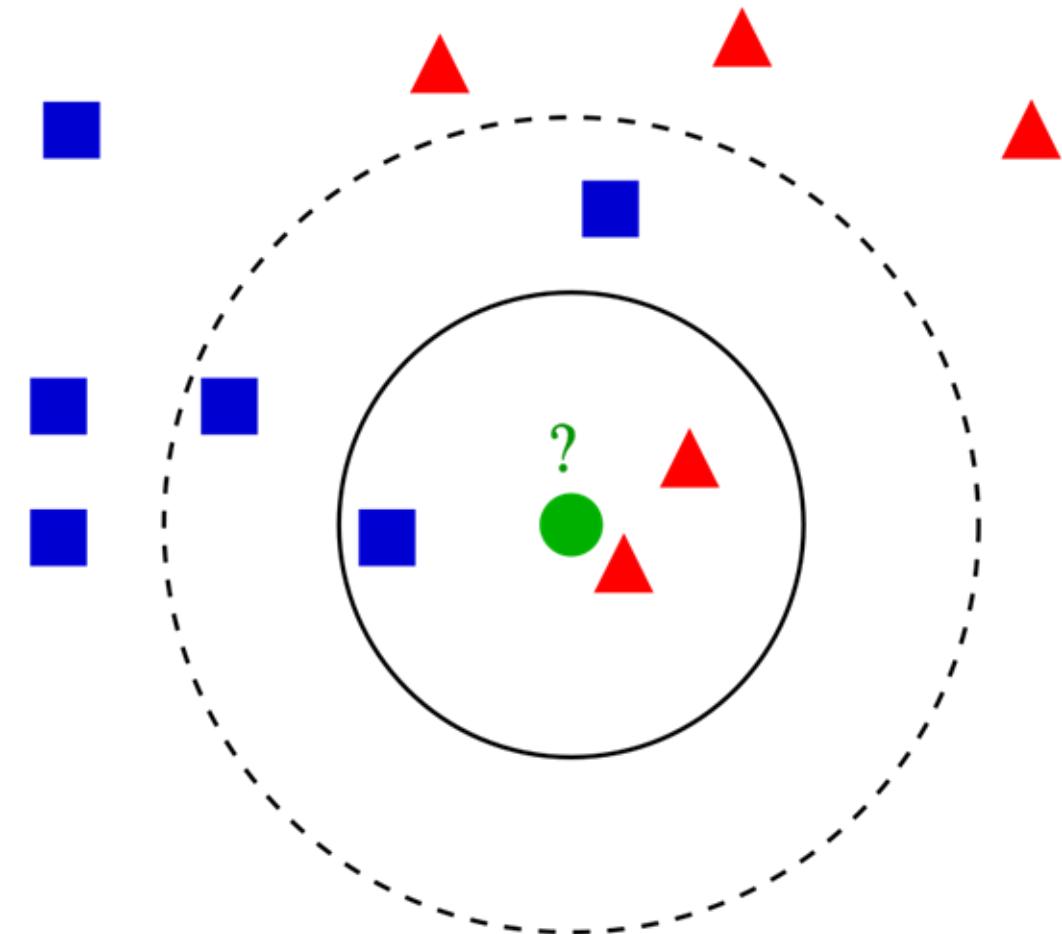
You typically split your dataset in 2 or 3 parts

1. Training set (typically 80%): Used for training the model
2. Validation set: Used for optimizing the model (optional)
3. Test set (typically 20%): Used for evaluating the model



Algorithm: K-nearest neighbors

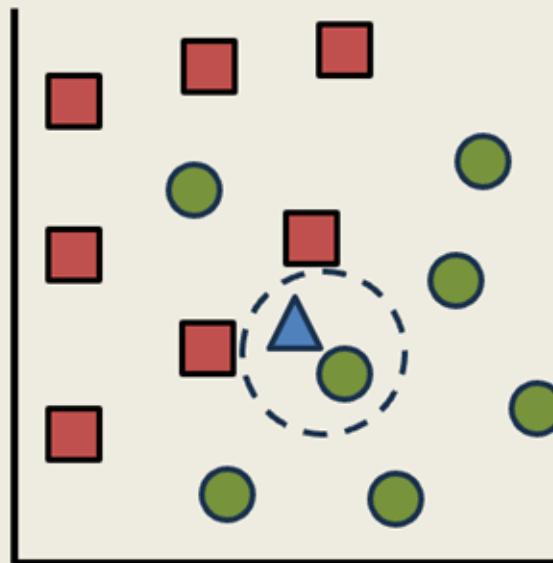
- Finds the k most similar observations in the training set and uses this as a basis for prediction.
- For a new observation you find the k most similar observations and **take the most occurring category**.



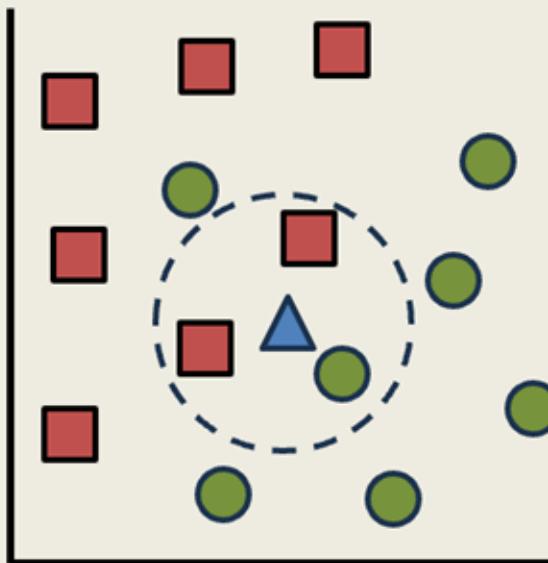
Algorithm: K-nearest neighbors

■ = Customer left

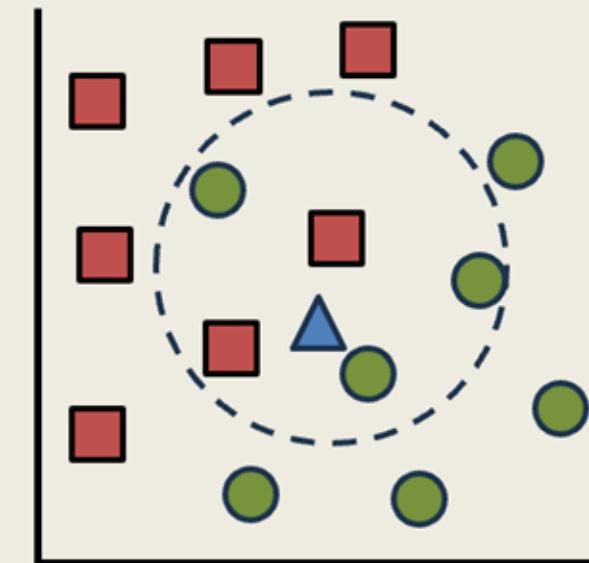
● = Customer stayed



$k = 1 \rightarrow \triangle = ?$



$k = 3 \rightarrow \triangle = ?$



$k = 5 \rightarrow \triangle = ?$

Exercise: K-nearest neighbors

Using pen and paper, draw up the data as follows:

- Number of products sold on the horizontal axis,
- Number of years as a customer on the vertical axis,
- Use 😊 for a customer who stayed,
- Use 😞 for a customer who left

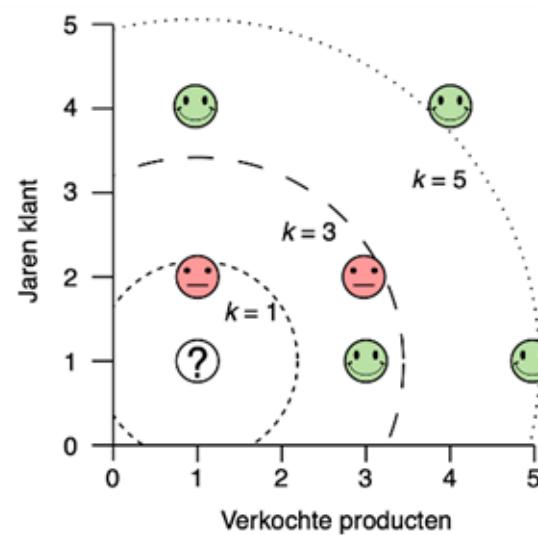
What is your prediction with $k = 1, 3$ and 5 when products sold = years as customer = 1 ? 😊 😞 😞 ?

What is your prediction with $k = 1, 3$ and 5 when products sold = 5 and years customer = 3 😊 😞 😞 ?

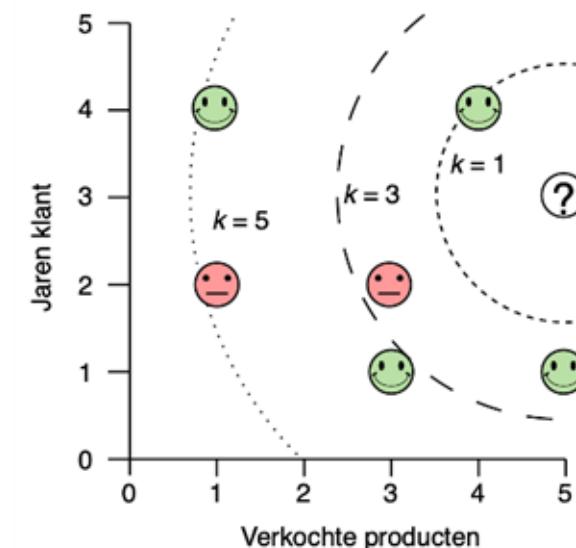
Sold products	Years customer	Outcome
1	4	Stayed
4	4	Stayed
1	2	Left
3	2	Left
3	1	Stayed
5	1	Stayed

Exercise: K-nearest neighbors

- What is your prediction with $k = 1, 3$ and 5 when products sold = years as customer = 1 ? 😊 😕 or 😕 ? 😊
or 😕 ?
- What is your prediction with $k = 1, 3$ and 5 when products sold = 5 and years customer = 3 😊 or 😕 ?



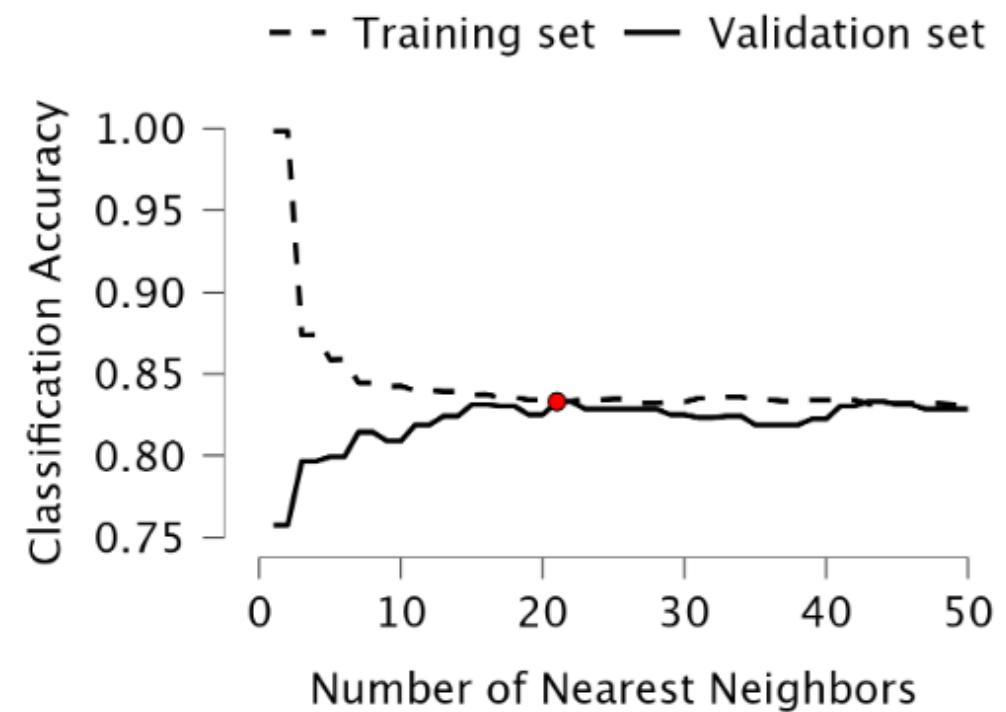
$k = 1 \rightarrow$ 😕
 $k = 3 \rightarrow$ 😕
 $k = 5 \rightarrow$ 😊



$k = 1 \rightarrow$ 😊
 $k = 3 \rightarrow$ 😊
 $k = 5 \rightarrow$ 😊

Validation: K-nearest neighbors

- $k = 1$ is a more flexible model, it gives perfect predictions on the training set, but does not generalize well to never before seen data
- $k = n$ is a more generalized model, it gives poor predictions on the training set but generalizes better to never before seen data.
- Optimal: somewhere in between.





| JASP

Exercises!



edu.nl/8thy7

Thank you for your attention!

Library Academy 2026

