Chart Advisor

Efficient Algorithm for Recommendation of Data Visualization Tools

Group Members
Cristobal Leiva
Ahmad Amayri
Jorge Ortiz

Supervisor
Fabrizio Orlandi

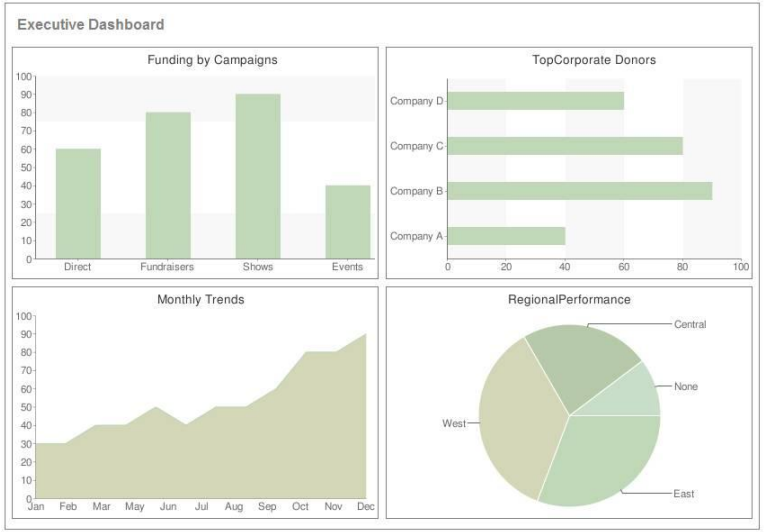
universitätbonn EIS ENTERPRISE INFORMATION SYSTEMS

System Overview

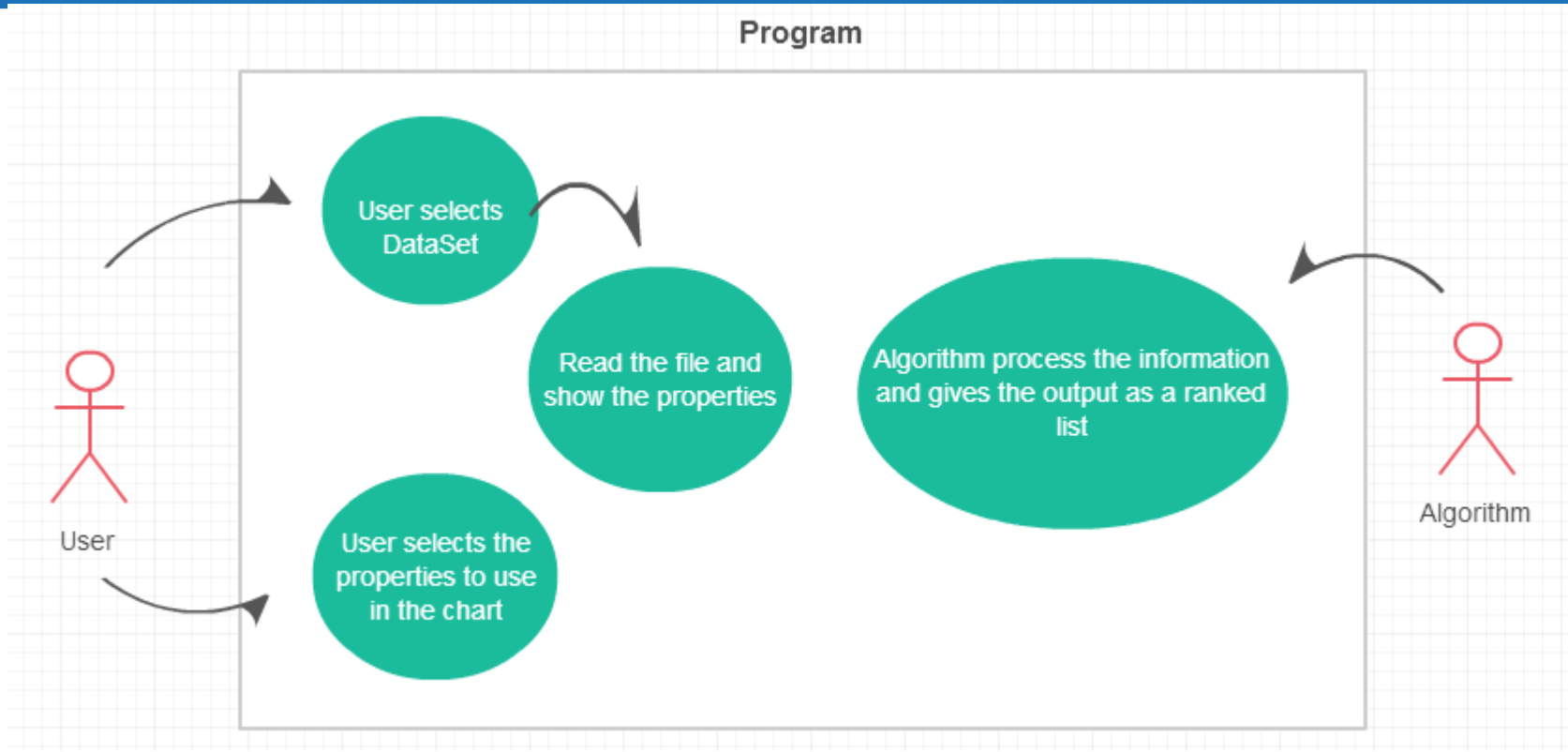
Project:

Develop an algorithm to recommend accured data visualization tools (Charts) based on selected Datasets / Properties

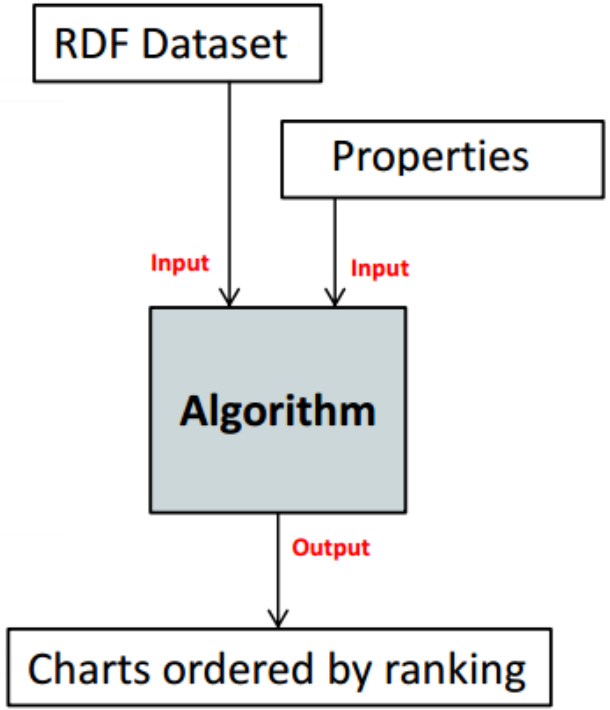
The output of the algorithm is a ranked list of recommended charts such as Bar, Bubble, Line, Geo Charts.



Use Case - General Scenario



Use Case - General Scenario



General steps in the algorithm

1. List and categorize properties.
2. Find out the pattern of the properties selected. (Quantitative, Qualitative or Ordinal).
3. Validation of data selected.
4. Suggest visualizations.

System Architecture

The software architecture is a model view controller.

Model.

Get the information from the Files and create the models with help of JENA libraries.

View.

Graphic interface based in SWING JAVA.

Controller.

Has the core algorithm.

HOW IT WORKS - ALGORITHM

Input:

- List of selected properties to visualize (User Selection)

Output:

- Recommended charts names and their accuracy in (RDF, XML, TXT)
  - Accuracy: determined by the number of attributes the suggested chart can visualize.

HOW IT WORKS - ALGORITHM

Input:

ID	Population	Region	Fertility Rate	Life Expectancy
DEU	81902307	Europe	1.36	79.84
CAN	33739900	North America	1.67	80.66
DNK	5523095	Europe	1.84	78.6
...	...	...	...	...

HOW IT WORKS - ALGORITHM

Categorize Properties:

- Determine the type and level of measurement
- Dictionary.rdf

Property	Type	Level of Measurement
ID	String	Categorical
Population	Number	Quantitative
Region	String	Categorical
Fertility Rate	Number	Quantitative
Life Expectancy	Number	Quantitative

HOW IT WORKS - ALGORITHM

Generate and Validate Allocations

- Allocations are the combinations of the input properties.
- Valid Allocation:
  - Left-Total: for any x in X, there is y in Y such that xRy
  - Right-Unique: xRy, xRz ⟶ y=z
- ID, Population ⟶ Fertility Rate, Life Expectancy
- ID, Population, Region, Life Expectancy ⟶ Fertility Rate
- ID ⟶ Fertility Rate
- ...

HOW IT WORKS - ALGORITHM

Map allocations to charts:

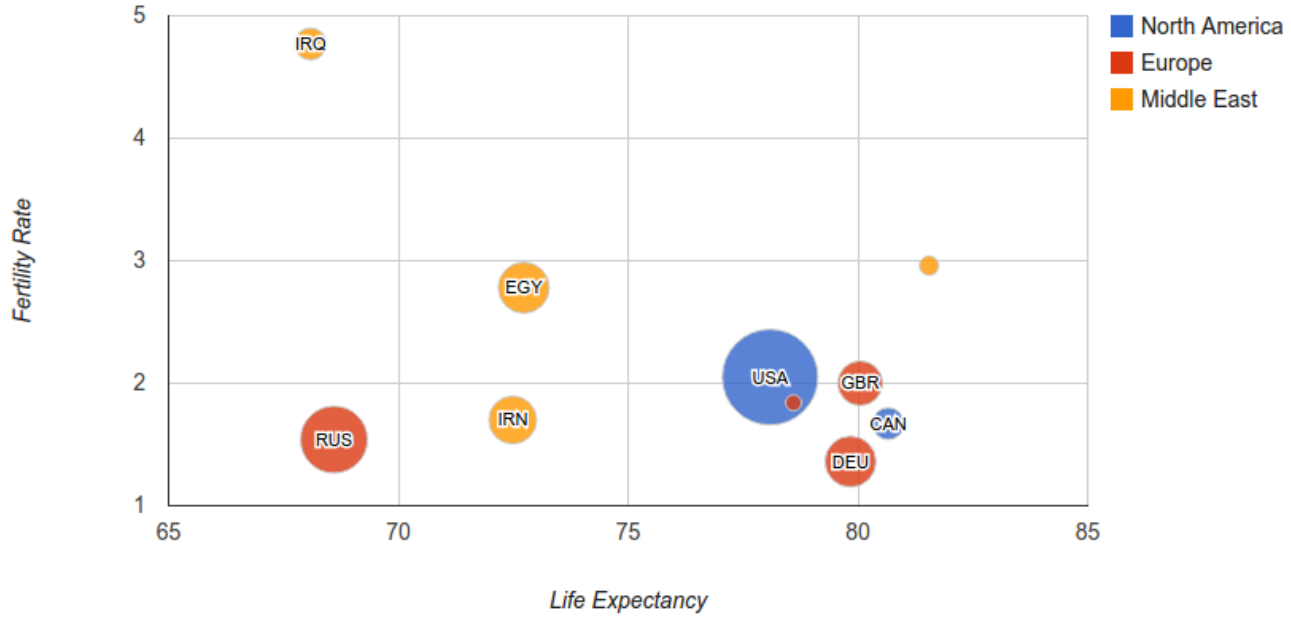
- Every chart capability is stored in chart.rdf
- Map valid allocations to existing charts
- Order by the number of mapped attributes.
  - Bubble Chart 100%
  - Bar, Column Charts 80%
  - Geo Chart 40%

Output:

- Save results to output file (RDF, XML, TXT)

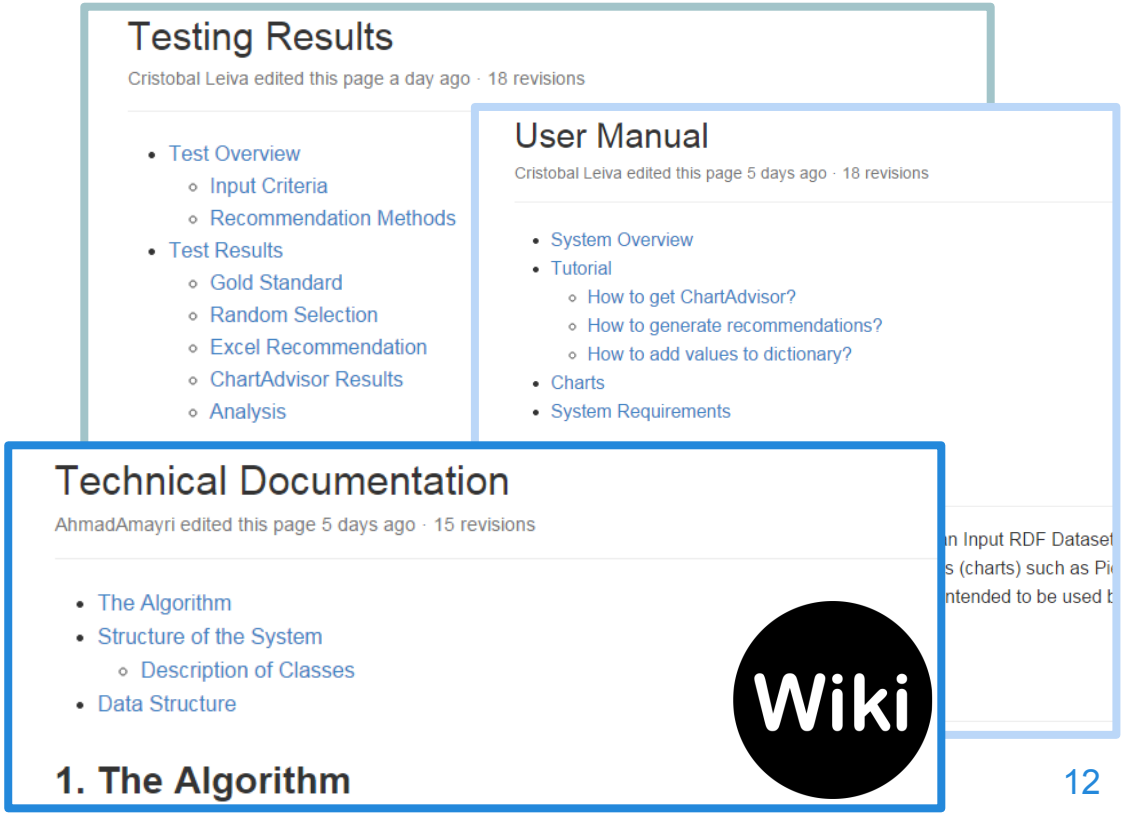
HOW IT WORKS - ALGORITHM

Bubble chart 100%



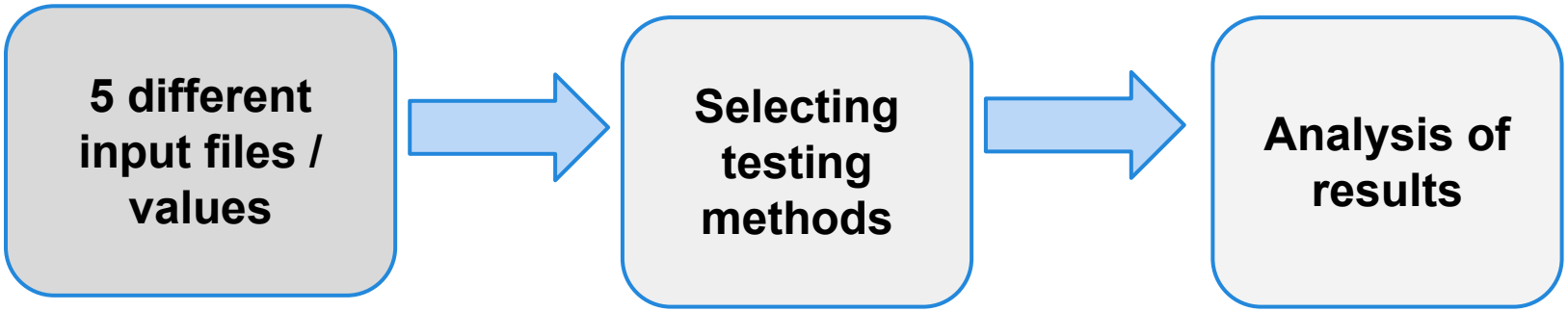
Documentation

- Technical Documentation
  - E.g. UML diagrams, architecture, algorithm explanation, more...
- User Manual
  - E.g. video tutorials, screencasting GIF images, more...
- Test Documentation
  - E.g. results tables, statistics, more...

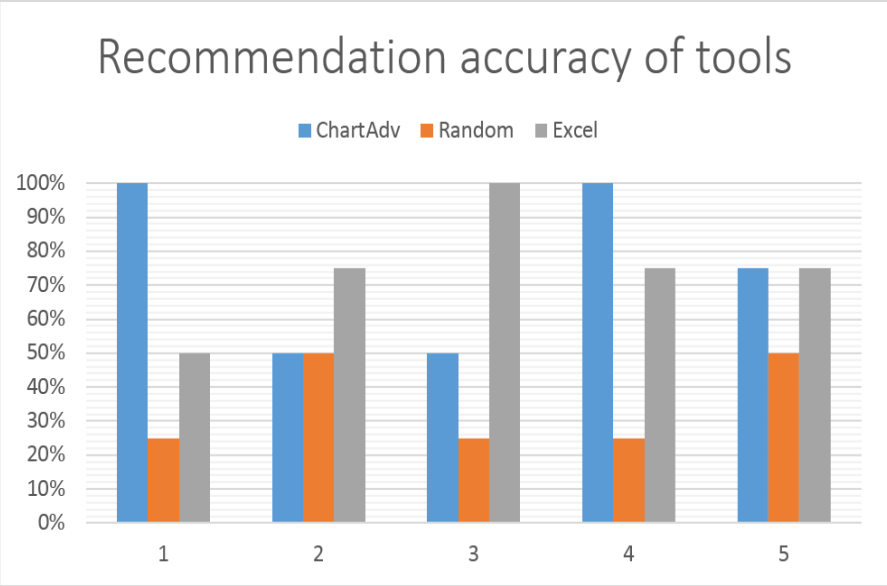


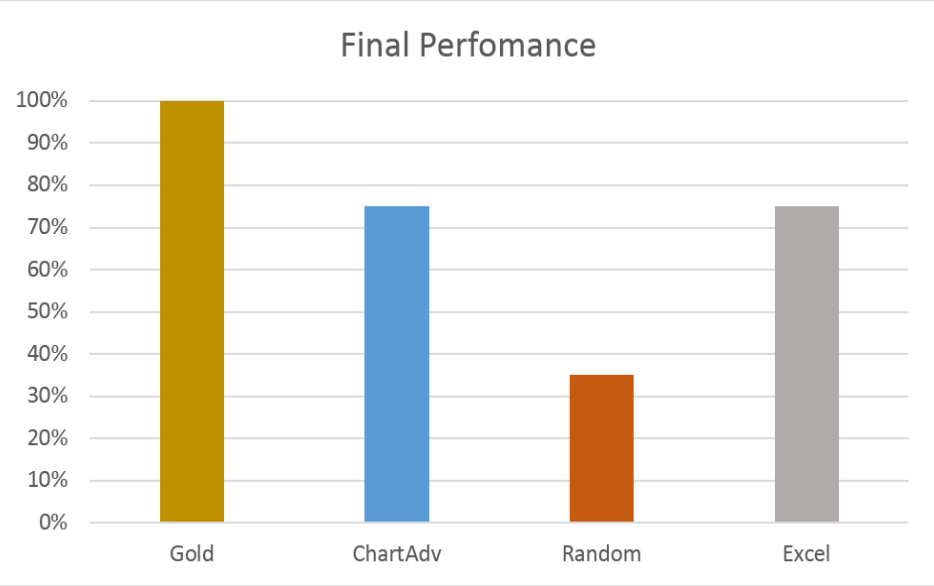
Testing

- Selection of input test samples: formats (.rdf and .ttl)
- Creation of Gold Standard charts recommendations.
  - Charts selected by team members.
- Testing methods.
  - MS Excel 2013 / Random selection / Chart Advisor
- Analysis of results



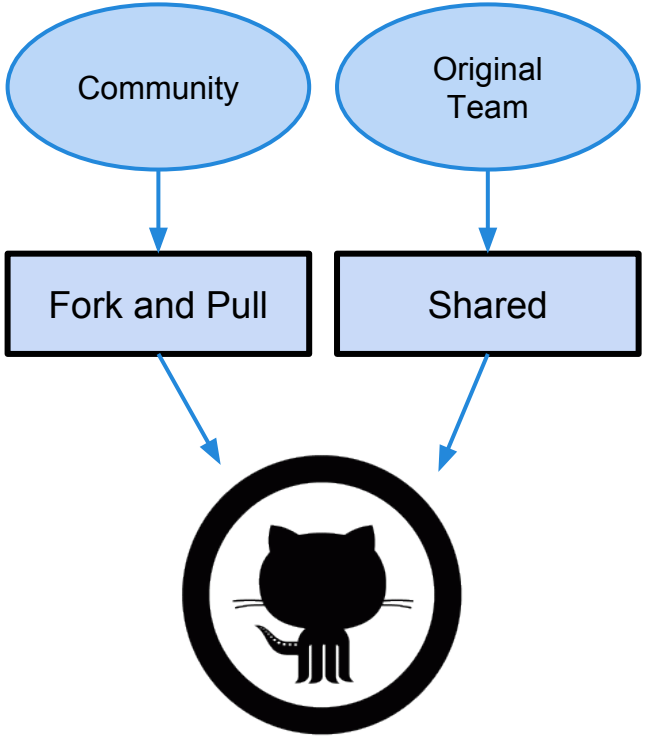
Testing - Results





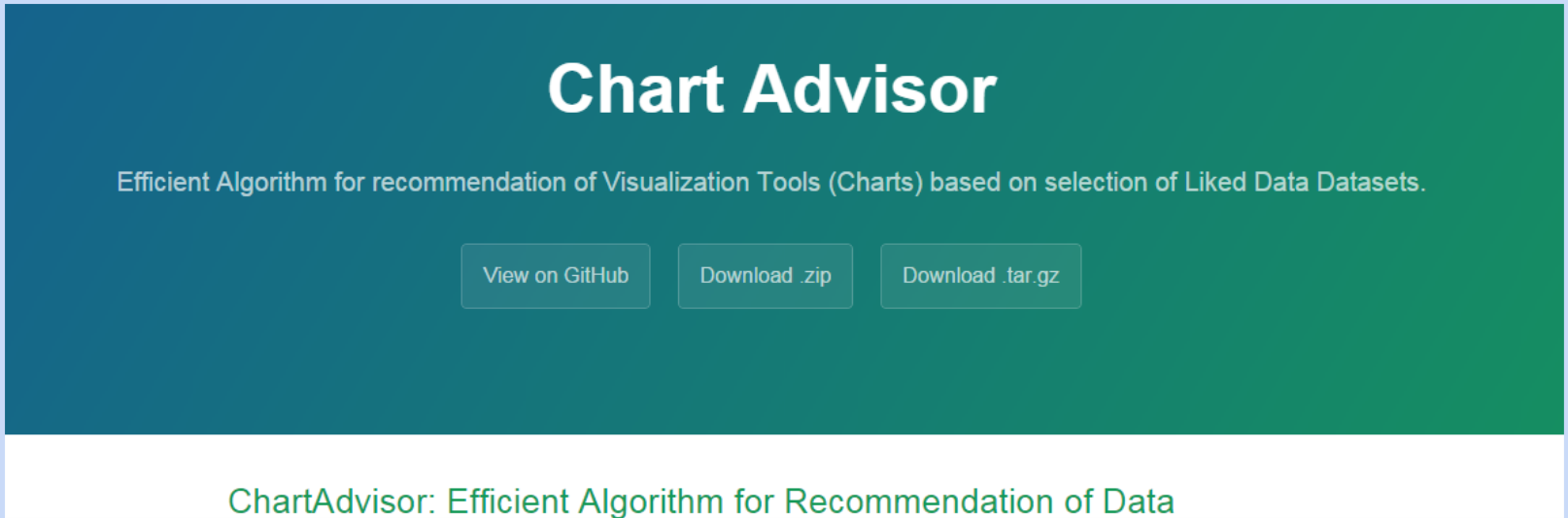
Deployment on GitHub

- Documentation full written in github wikis.
- Dedicated repository: Summary, source code, documentation.
- README.md with organization of the content on the repo. e.g. documentation index, libraries.



Project Homepage

- One-click download CartAdvisor and documentation.
- Better presentation of README content on website.



Show: Poster & Demo

Thank You

Group Members
Cristobal Leiva
Ahmad Amayri
Jorge Ortiz

Supervisor
Fabrizio Orlandi