MAKERERE  UNIVERSITY

SEMESTER ONE 2024/2025 ACADEMIC YEAR 1

SCHOOL COMPUTING AND IMFORMATICS TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

MASTER OF SCIENCE IN COMPUTER SCIENCE

MCN 7105

Structure & Interpretation of Computer Programs

Final Assignment

Eugene Munyaneza

2024/HD05/21936U

2400721936

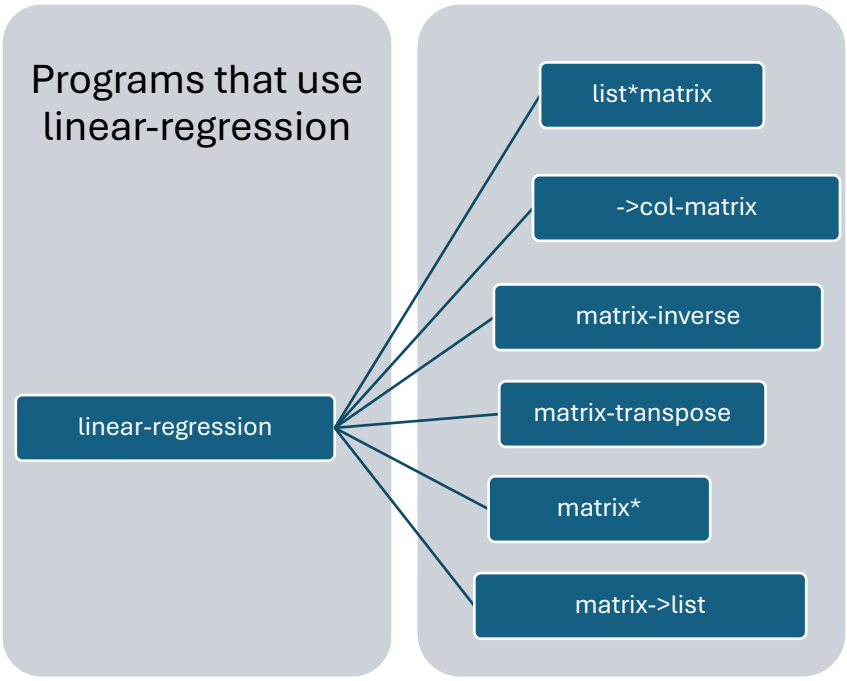
Introduction.

An abstraction diagram in computer science is a conceptual representation that illustrates the hierarchical relationships between different levels of abstraction in a program or system. It shows how high-level functions are built from lower-level ones and emphasizes the separation of concerns to simplify program design and reasoning.

1. Abstraction levels

a. linear-model

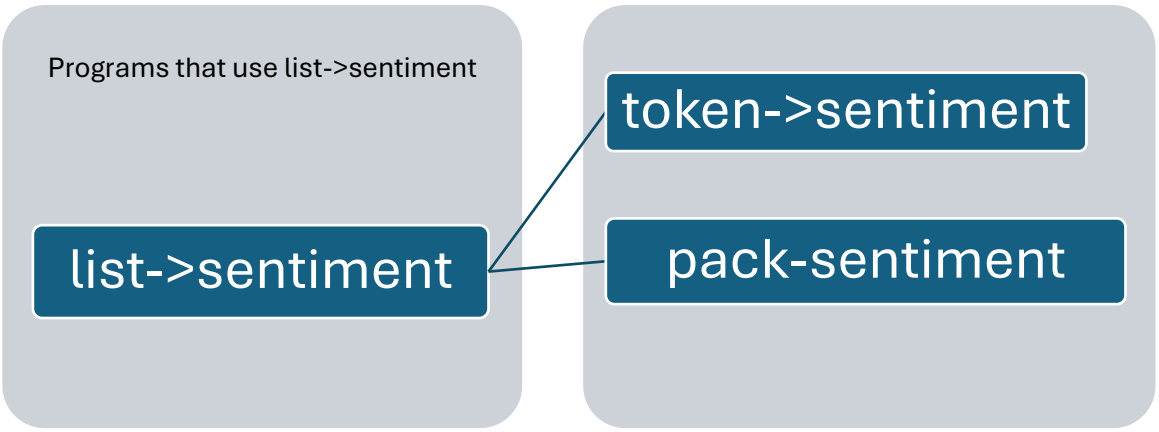
- Implements regression models, including linear regression.
- Code includes a slope, intercept calculation, and visualization using plot.
- Abstraction level:
    - **Data**: Input features (xs) and target variable (ys).
    - **Procedure**: Mathematical modeling and visualization.



b. list->sentiment

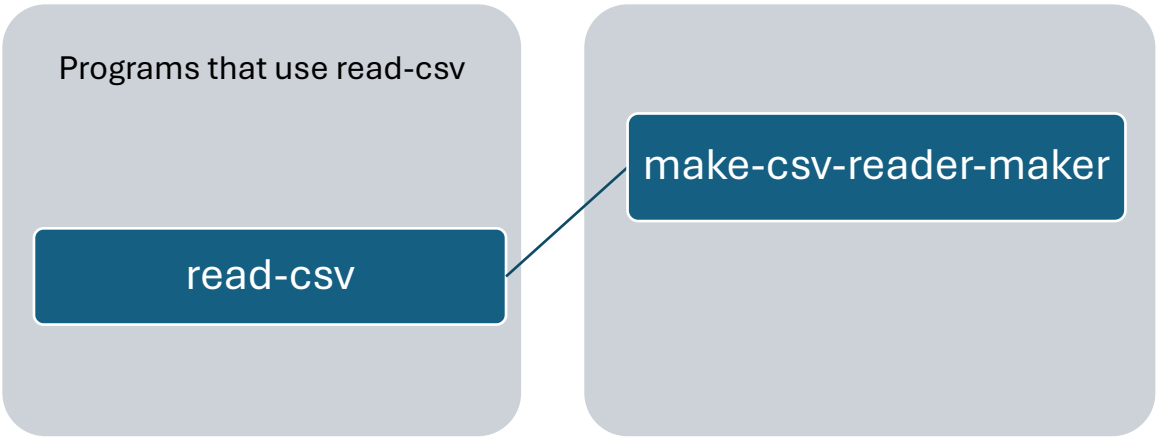
- Converts a list of tokens into sentiments using a specified lexicon (e.g., nrc).
- Abstraction level:
    - **Data**: List of tokens and their frequency.

- **Procedure**: Mapping tokens to sentiments and generating output structure.



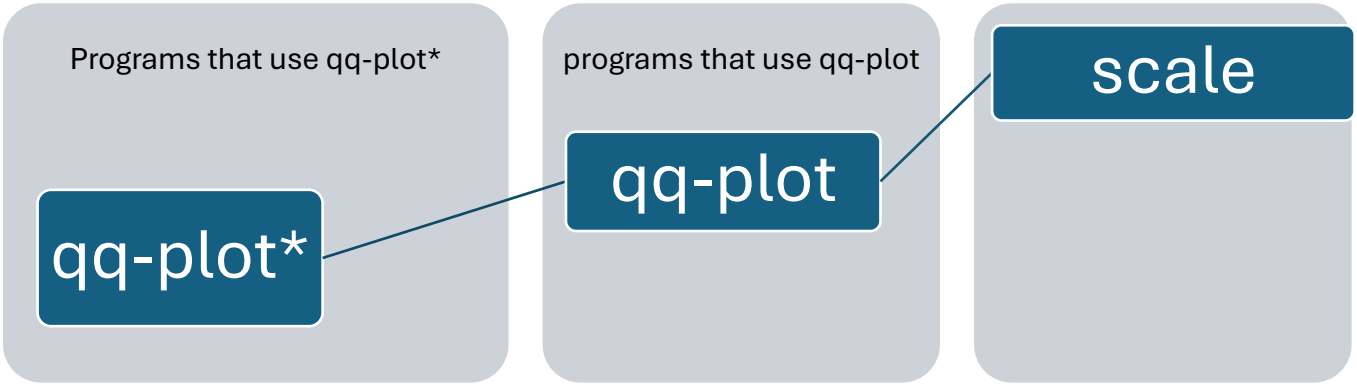
c. read-csv

- Reads CSV files, with options for numeric conversion and handling headers.
- Abstraction level:
  - **Data**: File input and structured output (list of rows).
  - **Procedure**: Parsing, type handling, and optional header processing.



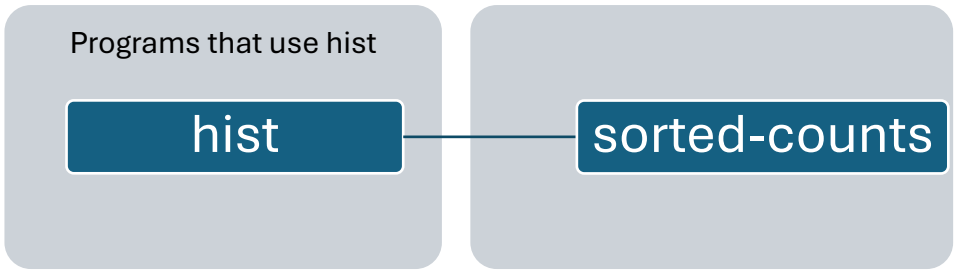
d. qq-plot*

- Generates quantile-quantile plots, with scaling options.
- Abstraction level:
  - **Data**: Input list (lst).
  - **Procedure**: Statistical transformations and visual plotting.



e. hist

- Generates discrete histograms from sorted binned samples.
- Abstraction level:
  - **Data**: List of values.
  - **Procedure**: Frequency calculation and binning.



2. Sentiment analysis of tweets by country.

The system is designed to process data from tweets in a csv file in the format of date/tweet/location. The program filters out the tweet data by location and analyses them for sentiments.

I used a local csv file since accessing tweets has been changed when twitter changed to X. But the code should run on any twiiter feed with prior data manipulation to produce the same input format.

Input data set

I. Time of tweet
II. Text of tweet
III. Location where a tweet was made from

Abstractions used

Procedures from the data-science module were utilized to carryout the analysis.

Data-science

I. read-csv
II. remove-punctuation
III. document->tokens
IV. list->sentiment

Sample analysis.

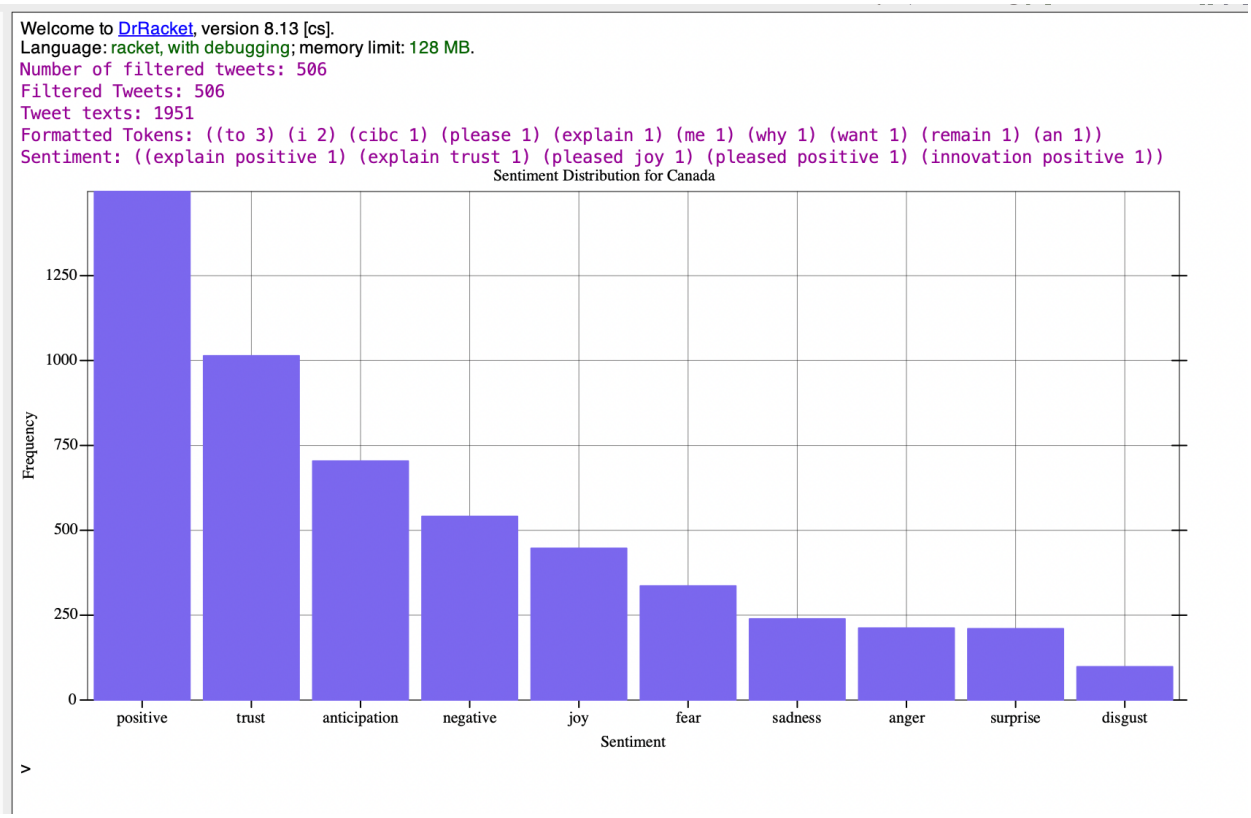


Figure 1 Sentiment analysis for Canada