**Homework 2**

PART 1: EXPLORE THE MAVEN DATASET (15 PT)

1. How many (unique) sentences are there in the documents? (2pt)

There are 14 unique sentences in the document.

1. How many (unique) event types are there in the documents? (2pt)

There are 27 unique event types in the document.

1. How many events of the type “Hostile\_encounter” occurred in doc2? (2pt)

There were 12 “Hostile\_encounter” in the document.

1. Which sentences mentioned the first Hostile\_encounter event (list the sentence ids)? (2pt)

Sentence ID: 0, 3, and 11 were mentioned by the first Hostile\_encounter.

Schema Design:

* Table: Battle of Salis
  + Attributes 1:
    - Name: Sentence
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 2:
    - Name: Tokens
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 3:
    - Name: id
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
* Table: Events
  + Attributes 1:
    - Name: id
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 2:
    - Name: type
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 3:
    - Name: type\_id
    - Data Type: INT
    - Constraints: NOT NULL
  + Attributes 4:
    - Name: trigger\_word
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 5:
    - Name: sent\_id
    - Data Type: INT
    - Constraints: NOT NULL
  + Attributes 6:
    - Name: offset
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 7:
    - Name: offset\_id
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
* Table: Negative\_triggers
  + Attributes 1:
    - Name: trigger\_word
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 2:
    - Name: sent\_id
    - Data Type: INT
    - Constraints: NOT NULL
  + Attributes 3:
    - Name: offset
    - Data Type: VARCHAR2
    - Constraints: NOT NULL
  + Attributes 4:
    - Name: offset\_id
    - Data Type: VARCHAR2
    - Constraints: NOT NULL

I would import the data from valid.jsonl to my table by separating the code by tables and attributes as I designed above, and would use DML to INSERT, UPDATE, and DELETE to create tables and attributes in an Oracle SQL developer environment. For each attributes, I would just copy and paste the corresponding attributes into the SQL database. Or for more efficiency, I could use the “json\_table” function to convert the JSON file into individual oracle attributes and tables.