# PROJECT CHECKPOINT

SEMANTIC WEB (S17)

**Project Title:** Heart or Tobacco – It's your choice!

**Team Name:** Mavericks

**Team Members:** Muhammad A Malik (mxm162431),   
 Teja Kiran Chunduri (txc163430)

**Type of Project:** Custom (changed from Simple because of no SPARQL endpoint)

**Progress:**

* Successfully set up SPARQL end point by setting up Fuseki server on our local machine.
* Completed LOGD tutorials that would help us through the project.
* Currently, exploring ways to convert our CSV data to RDF data.
* Explored many tools that claimed to convert CSV data to RDF data, but none of them gave satisfactory results so far.

**Completed Tutorials:**

* Building LOGD Visualizations
* Mashing up LOGD data with SPARQL
* Retrieving SPARQL Results
* Understanding LOGD Data
* Understanding LOGD Metadata
* Using MIT SIMILE Exhibit
* Getting started with csv2rdf4lod to create verbatim RDF conversions of tabular data

**Progress vs Planned Milestones:**

We are up to date with our planned schedule. Currently, we are working on converting our CSV data to RDF data using csv2rdf4lod which is a UNIX tools to convert CSV files to RDF files by following LOGD tutorials.

**Results:**

Querying over the 2 different datasets using their common attribute ‘locationabbr’ which is the abbreviation for states, the useful data values were returned as expected.

**Problems Encountered:**

Initially, our data from the data.gov files consisted of both CSV and RDF files, so we started directly by using RDF files. Querying over those RDF files was successful and gave expected results. As we explored more of those RDF files, we discovered that the RDF files provided in the data.gov site were only a part of the complete csv raw data and therefore one more task was added to our project timeline.

So, far we have explored most of the tools recommended on the LOGD website. The following tools gave unsatisfactory results:

* RDF123
* RDF Extension for Google Refine
* Sheet2RDF
* Spread2RDF
* XLWrap
* Tarql

**Changes:**

* No changes in the data sources
* No changes in the expected results