Project Proposal – Visualization of Yelp Check-In Data.

- **Background and Motivation.**

Our main motivations behind the project are threefold – one it gives us an opportunity to apply visualization techniques learnt in the classroom to spatial and temporal generated by everyday Yelp users, secondly the project would give us understanding of user behavior of checking into or visiting places based on various temporal patterns – time of day, day of week, month etc., thirdly it will also give us an opportunity to understand what different patterns of checking into events users exhibits based on demographics and age.

- **Project Objectives**
    - Learn and master spatial and temporal visualization techniques.
    - Understand user checkin behavior across demographics and time.
- **Data.**
    - We will be using the dataset from the yelp challenge dataset. The dataset can be obtained from here. We will primarily be looking at the check-in dataset.
- **Data Processing:** Although the yelp dataset looks clean on initial analysis, we envision that we will need to do some correlation of the data files in the yelp dataset. The current yelp dataset contains 4 files related to – users, tips, check-ins businesses and reviews. We will primarily looking at the business, users and check-in dataset and process them to a format useable in visualization.
- **Visualization.** How will you display your data? Provide some general

ideas that you have for the visualization design. Include sketches of your design.

    - The check-in data is will be displayed as heat map on the US maps.
        - The location would be encoded by position on map
        - The number of check-in at a time by color intensity.
    - Time-series to show the change of check-ins across time.
    - Adjoining charts show the heat map/ histogram of the user check-ins by time, user demographics etc..
- **Must-Have Features.** These are features without which you would consider your project to be a failure.
    - The check-in data is will be displayed as heat map on the US maps.
        - The location would be encoded by position on map
        - The number of check-in at a time by color intensity.
    - Time-series to show the change of check-ins across time.
- **Optional Features.** Those features which you consider would be nice to have, but not critical.
    - Adjoining charts show the heat map/ histogram of the user check-ins by time, user demographics etc..
    -
- **Project Schedule.** Make sure that you plan your work so that you can avoid a big rush right before the final project deadline, and delegate different modules and responsibilities among your team members. Write this in terms of weekly deadlines.
    - Week 1-2 : Planning and Drawing of Visualization sketch.
    - Week 3: Implementation of basic spatial data (checkins at an instance

of time) on map as heatmap.

- Week 4: Time series implementation of spatial data (checkins across the range of time) on the map as heatmap.
- Week 5-6 : Adjoining views such as histograms or area charts.