Coursera IBM Final Capstone Project

HOSPITAL COVERAGE FOR TRAFFIC COLLISIONS IN LOS ANGELES

Abhijit Deshpande

07-05-2020

Table of Contents

Project Introduction

Problem Background

- Victims of traffic accidents need to be quickly taken to the nearest hospital.
- Any individual hospital can provide timely health services only to accident victims located within a coverage radius of 10 to 12 km around it. The hospital coverage zone can be visualized as a circle with the hospital at the centre and radius as coverage radius, which is set at a distance of 12 km for the purpose of this analysis.
- **Having a hospital nearby the accident location can be the crucial difference between life and death.**
- Assuming that Los Angeles city administration wants to open a few new hospitals for catering to traffic collision victims, this analysis aims to identify the optimum locations in order minimize hospital coverage gaps.

Project Scope

- The scope of this analysis is limited to the city of Los Angeles (County level is out of scope).

Target Audience

- Local Administration of Los Angeles City is the primary target audience.
- Private players interested in opening new hospitals in Los Angeles are the secondary target audience.

Assumptions

- The coverage radius for an individual hospital is assumed to be 12 km for this analysis.
- It is assumed that future traffic collisions will follow similar pattern as last four years.
- Resources like funds, space along with necessary permissions are assumed to be available for opening few new hospitals at proposed locations.

Objective

Here are the objectives of this project along with the specific questions that are aimed to be answered:

- **Visualize last few year's traffic collision history in Los Angeles**
  - Do the central regions of the city have higher intensity of accidents than the outskirts?
- **Identify the hospital coverage gaps**
  - What percentage of recent traffic accidents occurred too far from the nearest hospital (outside hosp. coverage zones)?
- **Provide policy recommendations to Los Angeles administration**
  - If the LA city administration can afford to open a few new hospitals, where should they be located in order to minimize the hospital coverage gaps?
  - How many hospitals would need to be opened and where for reaching coverage target of
    - 85%?
    - 90%?
  - If the administration has resources to open 3 new hospitals, what coverage level can be achieved?

Data

What data was used?

- Location data of last few year's traffic collisions in Los Angeles city (2016 onwards)
  - Was downloaded from the public dataset at this link: https://data.lacity.org/api/views/d5tf-ez2w/rows.csv?accessType=DOWNLOAD
  - This public dataset was uploaded to IBM Cloud Storage for being used by code

- Location of hospitals in Los Angeles city: Was obtained through a combination of
  - Foursquare API: Details described in the "Data Processing Methodology" section
  - Public Dataset:
    - The public dataset of hospitals was downloaded from the this link: https://hub.arcgis.com/datasets/lacounty::hospitals-and-medical-centers
    - This public dataset was uploaded to IBM Cloud Storage for being used by code

Data Processing Methodology

- Traffic Collisions
  - Dropped rows: Older records up to year 2015
  - Dropped columns: Most columns like accident date, victim details along with rows outside LA city limits
  - Feature Extraction: Latitude and Longitude were extracted to separate columns from a single column in the raw data

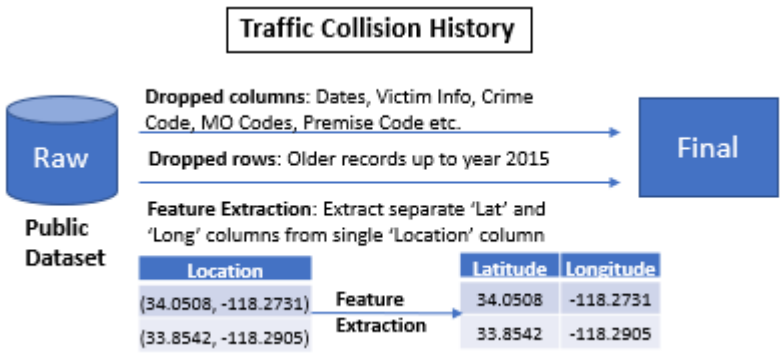


Figure 1: Traffic Collision Methodology

- Hospitals
  - Foursquare API Search

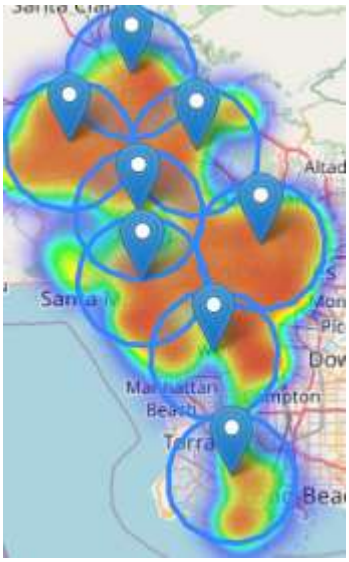


Figure 2: Hospital Search Zones for API Search

- Foursquare API limits the results to 50 results per zone. Many of these results were duplicates.
- Instead of using a single search zone with a huge radius, the city was divided into 8 search zones using trial & error and visual inspection. This helped in fetching more results.
- The results from all the zone were aggregated into a hospital master table.
- Duplicates with the same hospital names were dropped.

  - Public Dataset
    - The public dataset lists hospitals from all cities in LA county, not just LA City.
    - The latitude and longitude of the public dataset were filtered with upper and lower limits. This limited results to a bounded rectangle encompassing LA City.
    - The filtered results were merged with the master list of hospitals obtained using Foursquare API search
    - Duplicates with the same hospital names were dropped to get the final table.

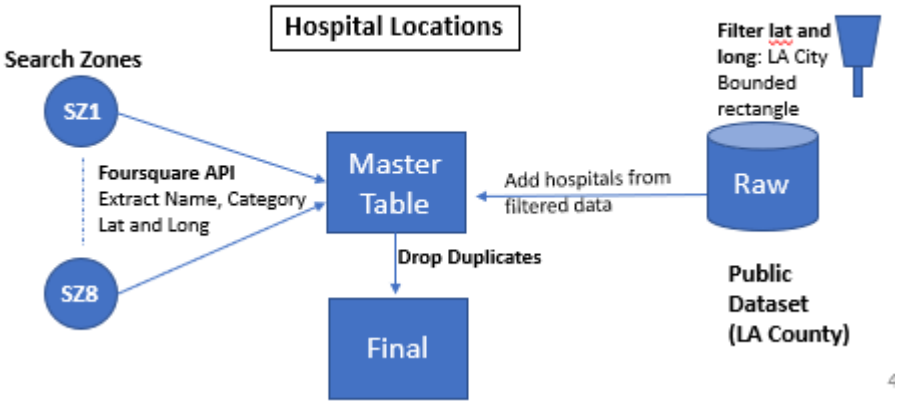


Figure 3: Hospital Locations Methodology

Sample Final Data for Processing

Traffic Collision History

	Area Name	Address	Cross Street	Latitude	Longitude
1	Rampart	UNION	JAMES M WOOD	34.0508	-118.273
2	Harbor	VERMONT	KNOX	33.8542	-118.291
3	Hollywood	FRANKLIN AV	ARGYLE	34.1052	-118.325
4	Southwest	LA BREA AV	JEFFERSON BL	34.0255	-118.3548

Hospital Locations

	Name	Category	Lat	Long
1	Ronald Reagan UCLA	Hospital	34.0665	-118.4465
2	Healthline Medical Group / Urgent Care	Urgent Care Center	34.1945	-118.4642
3	Santa Monica UCLA	Hospital	34.0272	-118.4865
4	Sherman Oaks Hospital	Hospital	34.1603	-118.4495

Methodology

Technical Aspects

- Machine learning algorithms like regression and classification are not required for solving this problem.
- The problem was solved using a combination of visual inspection and quantification of coverage/coverage gaps.
- Folium library was used for map visualization and extracting coordinates of any location from the map.
- The objects used for quantification include GeoDataFrames and MultiPolygon objects which were applied to the historical traffic collision data superimposed on the hospital locations.
- Ensured appropriate coordinate reference systems were set to correctly calculate location within coverage zone versus visualization on the map.

Steps Followed

- As Is
    - Traffic collision intensity was visualized using a Heat Map.
    - Hospital locations were superimposed on this map.
    - Rough patches were identified with high collision intensity where hospitals were not present.

- Gap Analysis
    - The coverage circles of all hospitals were visualized on a map.
    - The Heat Map of traffic collision intensity was visualized only for accident locations lying outside the coverage of all the existing hospitals.
    - This led to a more accurate representation of the gaps on the map.
    - The coverage and coverage gaps were quantified using the historical collision data.

- To be (Proposed Solution)
    - Visual inspection of the existing gaps was used to determine new hospital locations.
    - The new coverage gaps, after implementing the proposal, were visualized and quantified.
    - Recommendations were provided based on the above analysis.

Results

As-Is

The following figure is the visualization of hospital locations superimposed on traffic collision intensity heat map.

Blue dots denote hospital locations and the heatmap shows the intensity of traffic collisions. Colours in increasing order of intensity are purple, light blue, green, yellow and orange. Thus, orange areas represent the areas with highest intensity of traffic collisions and purple denotes the areas with least collision intensity.

Patches of high intensity collision regions without hospitals were identified and represented by ovals on the map. These patches represent only rough regions of gaps as coverage circles of hospitals are not incorporated in the As-Is phase.

It can be seen that collision intensity tends to higher in the central regions as compared to the outskirts.

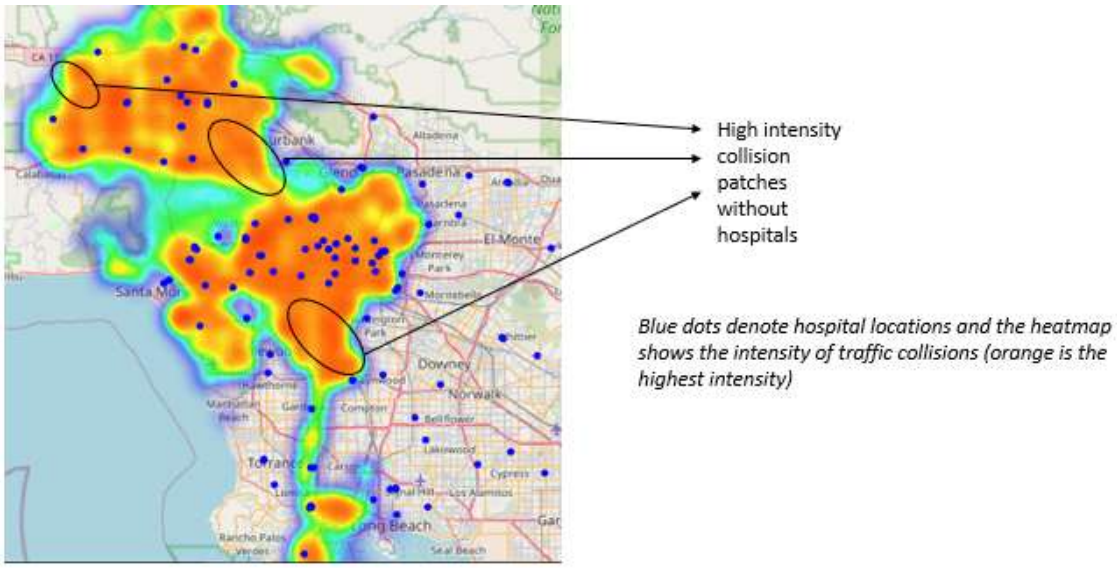
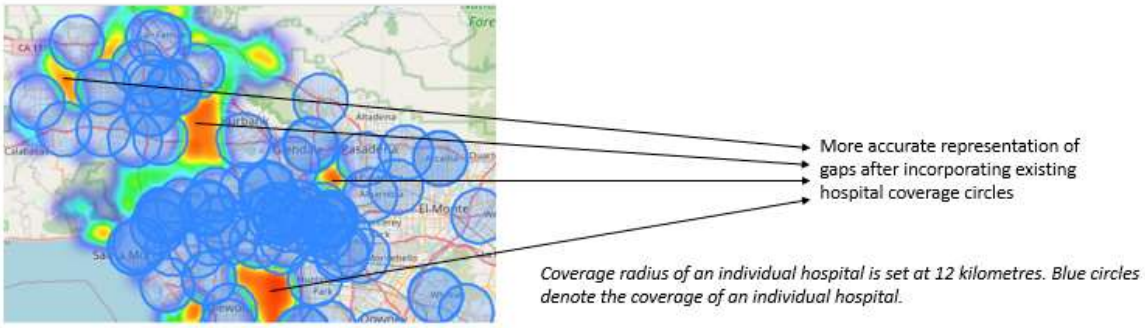


Figure 4: As Is

Gap Analysis

The following figure is the visualization of the gap analysis. The coverage circles of all hospitals were visualized on a map with blue circles. The Heat Map of traffic collision intensity was visualized only for accident locations lying outside the coverage of all the existing hospitals. This led to a more accurate representation of the gaps on the map.

The coverage and coverage gaps were quantified using the historical collision data. **Existing coverage was found to be 81% and consequently existing coverage gap = 19%.** Thus, 19% of traffic accidents in the last four years occurred more than 12 kilometres away from the closest hospital.

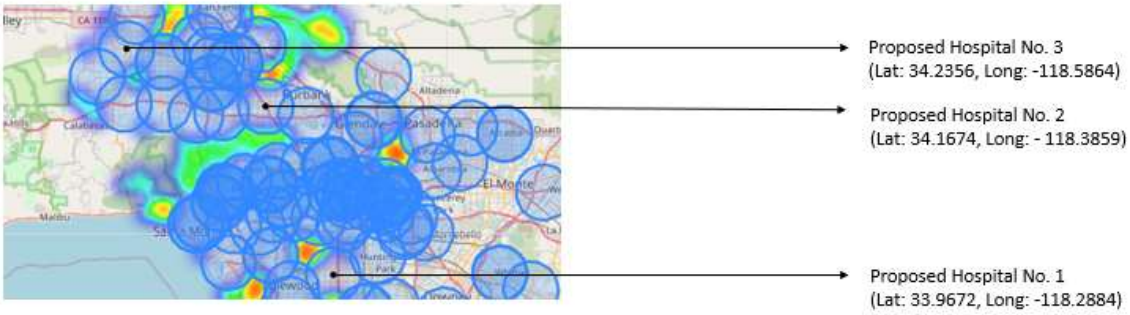


Existing Hospital Coverage = 81% and Coverage Gap = 19%
In other words, 19% of traffic accidents occurred more than 12 kms away from any existing hospital.

Figure 5: Gap Analysis

To Be

Visual inspection of the existing gaps from the Gap Analysis phase was used to determine new hospital locations. The new coverage gaps, after implementing the proposal, were visualized and quantified. Had the proposed hospitals been in place during the last four years, the coverage would have increased from 81% to 93%. Thus, coverage gaps would have been 7%. Assuming future traffic collisions will also follow a similar pattern as last four years, **we can expect that adding these 3 hospitals will bring up the hospital coverage to 93%.**



Adding the 3 proposed hospitals increases the coverage from 81% to 93%
(assuming that future traffic collisions follow similar pattern as the last 4 years)

Figure 6: To Be

Discussion

What is the contribution of each proposed hospital to the overall coverage?

The coverage benefits by adding each individual proposed hospital were calculated separately. This calculation was done by finding the reduction in percentage of total traffic collisions achieved by adding each individual proposed hospital against the historical traffic collision data.

The following figure shows the results. It can be seen that we see diminishing additional coverage benefit by adding more new hospitals. By adding the three proposed hospitals, cumulative coverage increased from 81% to 93%.

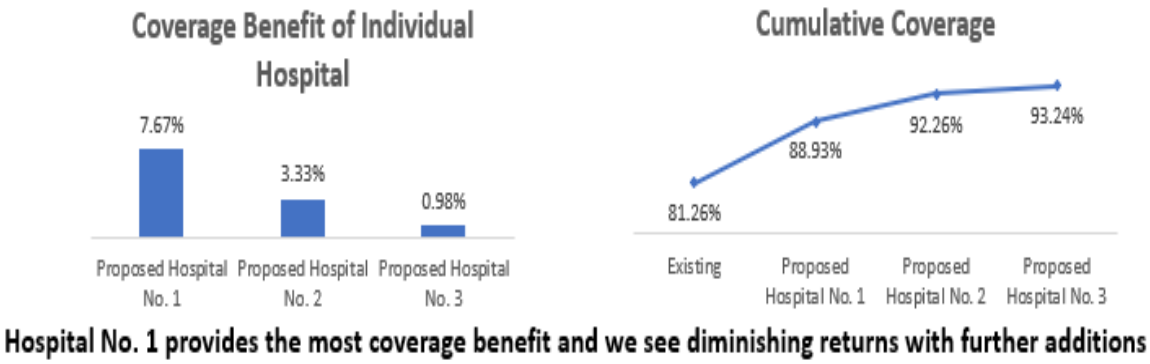


Figure 7: Contribution of proposed hospitals

Funding Options

- Proposed Hospital No. 1 and Hospital No. 2 are located in relatively central areas of the city.
- The health needs of residents nearby to Proposed Hospital No. 1 and No. 2 are unfulfilled currently, as there are no hospitals nearby.
- Both these hospitals are likely to be commercially viable for private players, as they can cater to all the basic health needs of nearby residents (not just limited to accident victims).
- Proposed Hospital No. 3 is situated at the outskirts, so may not be commercially viable for private players.
- The local administration can choose to fund a smaller scale hospital for Proposed Hospital No. 3 either in partnership with a private player or fully public owned, subject to availability of resources.

Recommendations

Here are the recommendations for various scenarios.

If the target is to achieve 85% coverage:

- **Building Proposed Hospital No. 1 is enough** to surpass the 85% coverage target.
- It can be built preferably through private players with city administration role limited to ensuring availability of space and permissions.
- In case private funding cannot be secured, local administration resources should be prioritized for this hospital.
- Proposed Hospital No. 1 can be full scale hospital catering to all basic health needs of nearby residents.

If the target is to achieve 90% coverage:

- **Build Proposed Hospital No. 1 and Proposed Hospital No. 2**, preferably funded and operated by private players.
- Both these can be full scale hospitals catering to all basic health needs of nearby residents
- In case private funding cannot be secured, local administration resources should be prioritized for proposed No. 1, then No. 2.

If 3 new hospitals can be built:

- **Coverage level of 93.24 % can be achieved** with 3 new hospitals.
- Proposed Hospitals No. 1 and 2 can be full scale hospitals, built preferably through private players.
- Proposed Hospital No. 3 can be smaller scale and may need to be built through public-private partnership or fully public owned.
- In case private funding cannot be secured, local administration resources should be prioritized according to the hospital number.

Achieving 95% target:

- Achieving a target of 95% would require many more hospitals added due to the diminishing incremental returns.
- This would not be possible to achieve with judicious use of public resources.
- Hence it is not recommended to add more than 3 hospitals.

Conclusion

- Considering recent historical collision data from 2016, existing hospitals offer a coverage for 81% of the collisions.
- Thus **19% of recent collisions occurred more than 12 km away from the nearest hospital.**
- Collision intensity tends to be higher in the central city regions as compared to the outskirts.
- Assuming that future traffic collisions will also follow a similar pattern as recent years - **Adding just one hospital can improve the coverage to 89%, adding two can lead to 92% and adding 3 can lead to 93% coverage.**
- The locations of the proposed hospitals are Proposed Hospital No. 1 (Lat: 33.9672, Long: -118.2884), Proposed Hospital No. 2 (Lat: 34.1674, Long: - 118.3859) and Proposed Hospital No. 3 (Lat: 34.2356, Long: -118.5864). Slight deviations from the target locations (within 500 metres) should still give similar results.
- New Hospitals must be **prioritized in the order of their individual coverage benefit**. Hence No. 1 has the top priority and No. 3 has last.
- The local administration must seek **private ownership model** to conserve its resources. **Proposed Hospitals No. 1 and No.2** are attractive candidates, being located in relatively central regions of the city and their services will not be limited to traffic collisions only.
- **Proposed Hospitals No. 1 and No. 2 can be full scale hospitals** catering to all the basic health needs of the nearby residents.
- Being located towards the outskirts, **Proposed Hospital No. 3 can be a smaller scale hospital** and can be funded through public – private partnership model or fully public owned.
- It is not recommended to add more than 3 hospitals in order to make judicious use of public resources.