

APPLIED DATA SCIENCE CAPSTONE

<https://www.coursera.org/learn/applied-data-science-capstone>

ASSIGNMENT: THE BATTLE OF NEIGHBORHOODS

“STARTING A NEW CINEMA IN ATHENS, GREECE

Dimitra Dionysiou

Data

The data sources that have been used in this analysis are the following:

- A list of the existing indoor cinemas in Athens, along with their addresses is retrieved from a well-known Athens city guide:
<https://www.athinorama.gr/cinema/guide.aspx?show=1&seltab=1&sec=2>
Each cinema of the retrieved list will serve as the center of candidate location for starting a new cinema.
- Geographical coordinate data for each cinema is retrieved from OpenStreetMap: <https://www.openstreetmap.org> using each cinema's address (Nominatim search engine, geopy Python package).
- Foursquare location data are used to decide which cinemas from the initially derived cinema list will be considered in the subsequent analysis, by excluding cases where more than 5 other cinemas exist in a 500m distance, according to the requirements defined in the Introduction section. The corresponding venue category id ("movie theater") is used for these API calls.
- Foursquare location data are also used to define the characteristics of each candidate location (centered on an already existing cinema) in terms of the features of interest: restaurants, nightlife spots, shopping malls and transport options (metro stations, bus stops, private parking places), according to the requirements defined in the Introduction section. The corresponding venue category ids are used to perform the necessary Foursquare API calls, by defining a radius of interest equal to 250m.