Making data-driven User Recommendations

Coursera Capstone

November 2020

What to do at a foreign train station?

- Assume your are at a foreign train station, you don't know the city, you don't have personal transport and you have limited time.
  - What can you do?
  - Where can you go?
  - Which places could you like, which not?
- We can use data to make recommendations.

What data do we need?

- List of local train stations we can take from our position.
  - Name
  - Location (addres, latitude, longitude)
- List of venues near train station.
  - Name
  - Location (addres, latitude, longitude)
  - Type (restaurant, bar, . . .)

What can we infer?

Distribution of different venue types in the city.

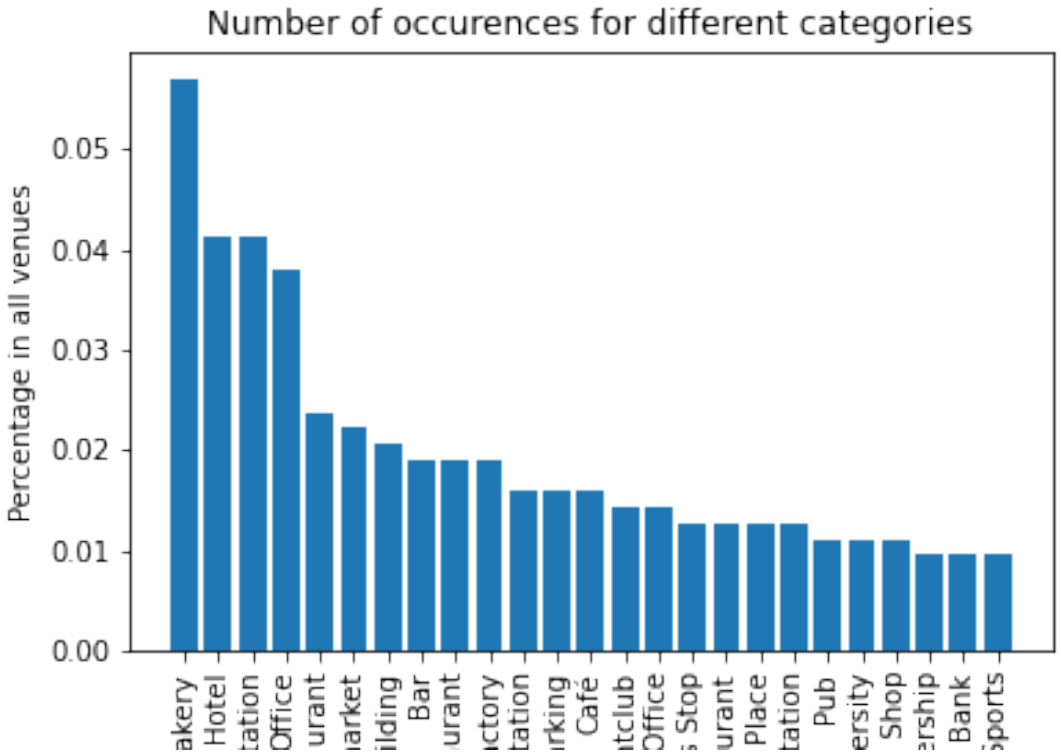

Figure: Top 25 venue categories in the dataset.

What do we do with this?

- We can group the train stations in groups with a similar venues nearby.
  - Easier to recommend venues.
  - Adjust recommendations to user.
  - Use machine learning.

Dividing the city in Clusters

We can plot these clusters:

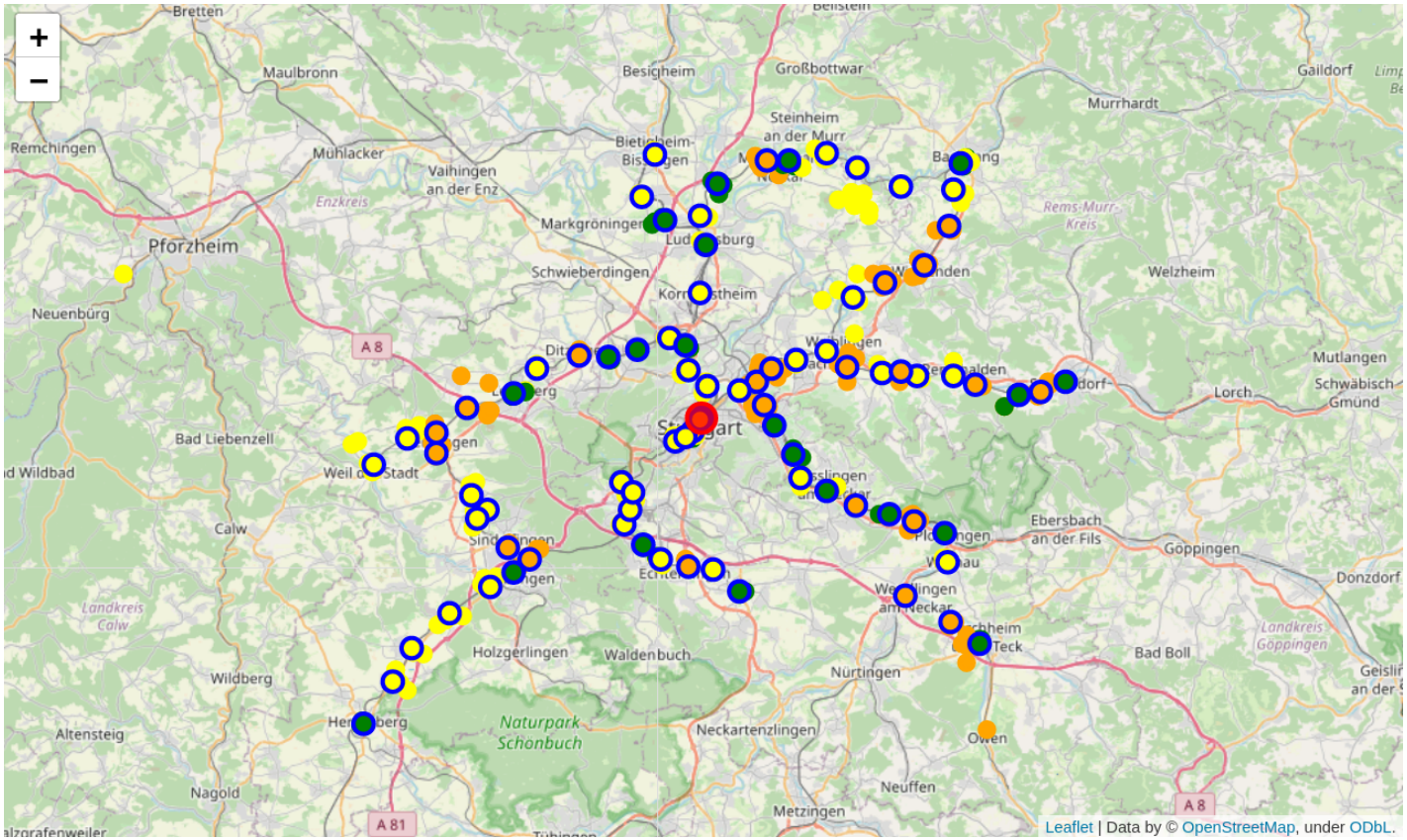

Figure: Plotted stations and venue, colorcoded by cluster. Unicolor circle are venues, circles with blue borders are station, red circle is exemplary user position.

Conclusion

- We saw what data we can use.
- We saw what inferences we could make.
- We saw how we can use this data.
  - We can make data-driven recommendation.
  - With a given target wish or just a general recommendation.

What else can we do?

- Add context information for better recommendations.
- Add other modes of transport.
- Search in a bigger radius and for more venues.
- Many possibilities to improve performance even more.