Introduction

The Berlin S-Bahn has 157 stations and is one of the largest public transit systems in Europe. It is used by 478 Millionen passengers per year.

For this project, we want to look at the neighborhoods surrounding S-Bahn stations and classify them. Some neighborhoods are mostly residential, others have more business or commercial spaces surrounding them. The venues closest to a station determine why and how people use it. E.g. if there are no professional places in a neighborhood its residents are likely to travel to other areas for work. This creates daily migrations of people.

By analyzing this data we can classify stations by their primary usage. This data can be useful for city planners to determine where from and where to people are most likely to travel for work and leisure, plan further extension of the network and find places for new development.

Data

We'll need data on the location of stations and on the venues closest to them.

1. List of stations and their geographical coordinates — scraped from this page:
https://de.wikipedia.org/wiki/Liste_der_Stationen_der_S-Bahn_Berlin

	Station Name	Line	Neighborhood	lon	lat
0	Adlershof(bis 1935 Adlershof=Alt-Glienicke)	Görlitzer Bahn	Adlershof(Treptow-Köpenick)52° 26′ 5″ N, 13° 3…	52.434722	13.541389
1	Ahrensfelde	Wriezener Bahn	Marzahn(Marzahn-Hellersdorf)52° 34′ 18″ N, 13°…	52.571667	13.565000
2	Albrechtshof	Hamburger Bahn	Staaken(Spandau)52° 32′ 58″ N, 13° 7′ 42″ O	52.549444	13.128333
3	Alexanderplatz	Stadtbahn	Mitte(Mitte)52° 31′ 17″ N, 13° 24′ 43″ O	52.521389	13.411944
4	Alt-Reinickendorf(bis 1994 Reinickendorf)	Kremmener Bahn	Reinickendorf(Reinickendorf)52° 34′ 40″ N, 13°…	52.577778	13.350556
5	Altglienicke	Grünauer Kreuz–BER	Altglienicke(Treptow-Köpenick)52° 24′ 26″ N, 1…	52.407222	13.558889
6	Anhalter Bahnhof	Anhalter BahnDresdener BahnNord-Süd-Tunnel	Kreuzberg(Friedrichshain-Kreuzberg)52° 30′ 11″…	52.503056	13.381944
7	Attilastraße(bis 1992 Mariendorf)	Dresdener Bahn	Tempelhof(Tempelhof-Schöneberg)52° 26′ 52″ N, …	52.447778	13.360833
8	Babelsberg(bis 1938 Nowawes)	Wannseebahn	Babelsberg(Potsdam)52° 23′ 29″ N, 13° 5′ 32″ O	52.391389	13.092222

Cut-Out of the table with the location data of the S-bahn stations, data already cleaned

2. Foursquare API to explore venue types surrounding each station. Foursquare outlines these high-level venue categories with more sub-categories.

- Arts & Entertainment (4d4b7104d754a06370d81259)
- College & University (4d4b7105d754a06372d81259)
- Event (4d4b7105d754a06373d81259)
- Food (4d4b7105d754a06374d81259)
- Nightlife Spot (4d4b7105d754a06376d81259)
- Outdoors & Recreation (4d4b7105d754a06377d81259)
- Professional & Other Places (4d4b7105d754a06375d81259)
- Residence (4e67e38e036454776db1fb3a)
- Shop & Service (4d4b7105d754a06378d81259)
- Travel & Transport (4d4b7105d754a06379d81259)

We'll be querying the number of venues in each category in a 1000m radius around each station. This radius was chosen because 1000m is a reasonable walking distance.