DATA ANALYSIS

Scrapping Yelp

**PARAMETERS**

* Location: Berlin
* Activities:
  + Restaurants
  + Pubs:
    - Bars
    - Nightlife-- Do we need to include this category?
  + Hotels: needs a specific search- we need to choose the categories considered such as:
    - Pensions
    - Hotels
    - B&B

**ABOUT DATA**

Goal: Give a final recommendation on the business type, where this business should be and key factors to make succeed, avoiding bad press through Data (Overall public opinion).

* Business info from YELP
  + Suggested by Slides
    - Rating
    - Location🡪 filtered by neighbourhood (we can try to check the accessibility)
    - Reviews 🡪 users with “Elite” tag; Ratio between rating and reviews
    - Amenities or Features
  + Other ideas
    - Name (Restaurants, hotel and pub)
    - Types of restaurants
    - Opening hours
    - Dine-in or take away
    - Price range (€€€)
    - Number of reviews
    - Reviews origin (we can have a general idea if most of the reviewers are locals or tourist)
    - Review date (post COVID)
  + For Hotel:
    - Photo gallery
    - Price range (€€€)

**HOW TO REACH THIS DATA?**