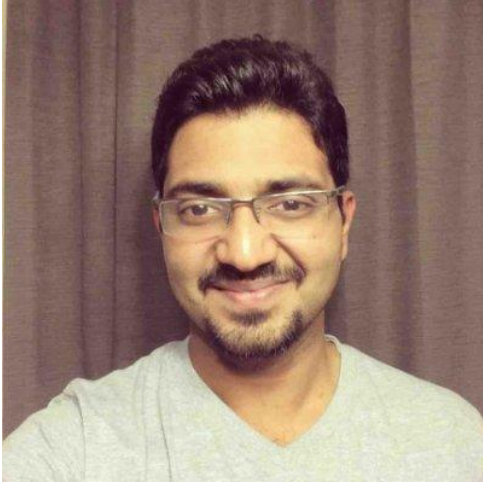


# **AirBnb – Home away from home**



# The Team



**Karthik Balasubramanian**  
Masters in Information Systems

**Shuting Zhang**  
Masters in Information Systems



**Ankit Goyal**  
Masters in Information Systems



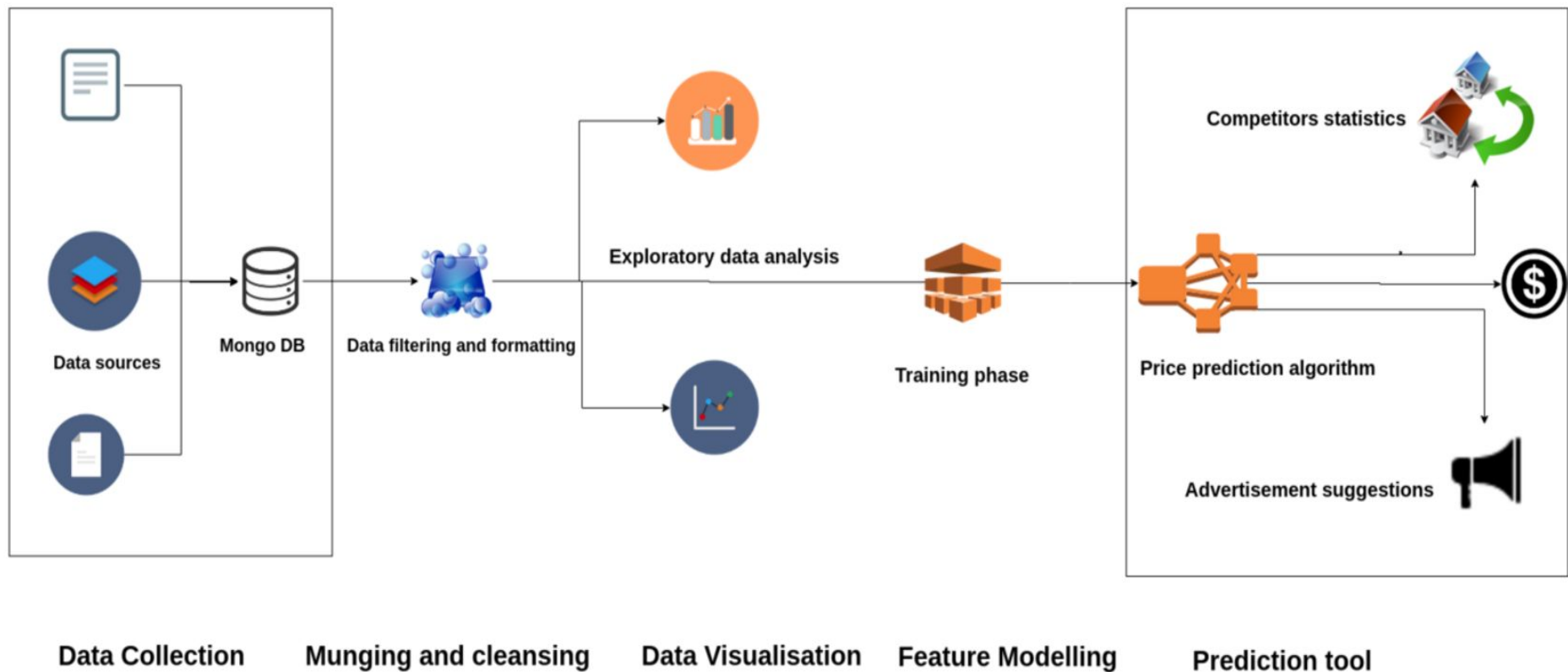
“

## **Problem Statement**

*What should be the price of an Airbnb listing given its features, popularity of its neighborhood and its competitors? How to advertise a listing to maximize its capacity?*

# Methodology

## Data Science Pipeline



# Data Collection and Munging

## © Airbnb API

Listing Search (listing id, ratings, reviews and etc.)

View Listing Info (listing location, amenities and etc.)

View User Info (guest name, country and etc.)

## © geocoder library

# Exploratory Data Analysis

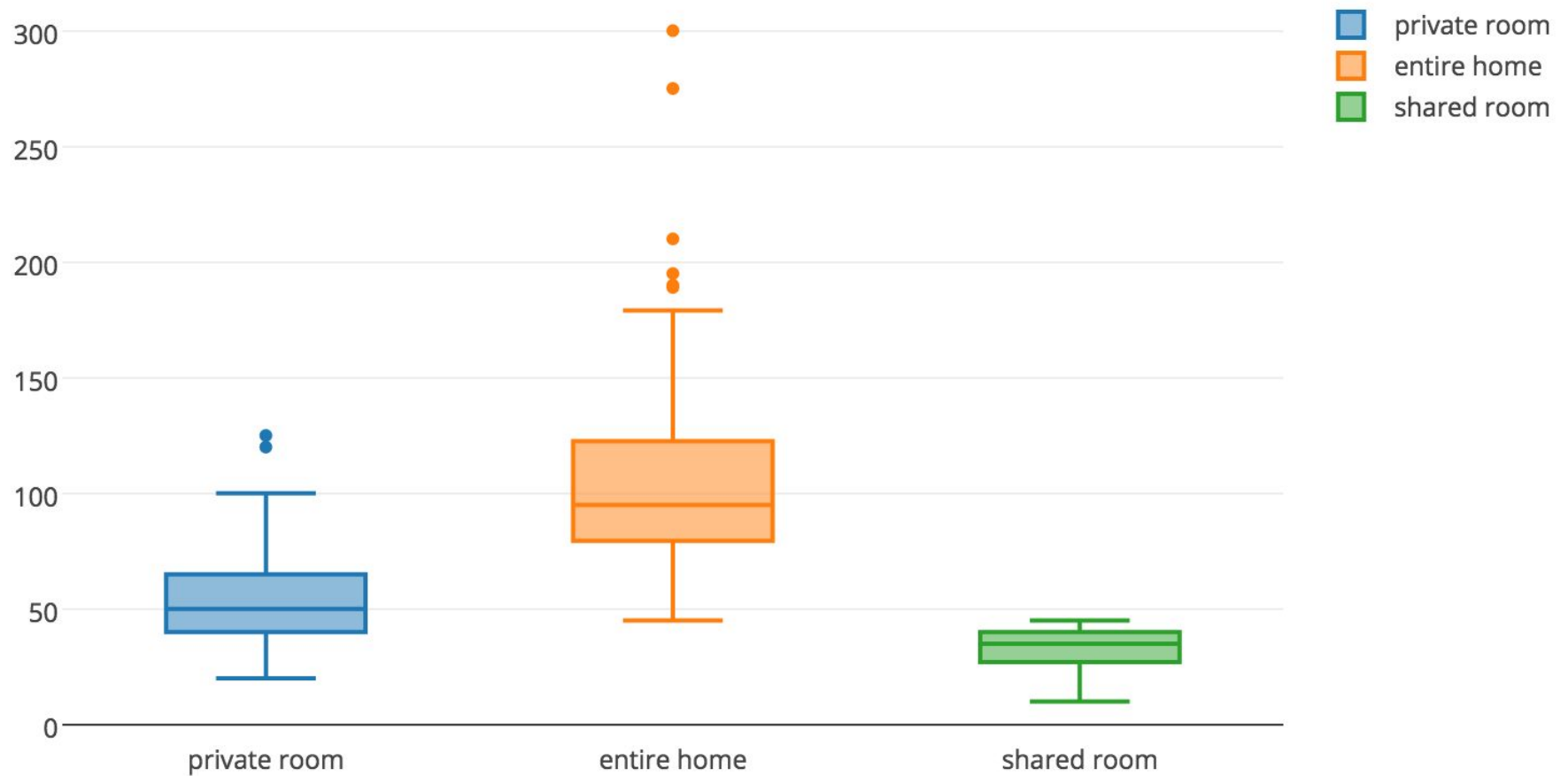
A decorative network diagram in the top right corner, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey, creating a mesh-like structure.

## Price

How does price change over different room types?

A decorative network diagram in the bottom left corner, similar to the one in the top right, featuring a complex web of interconnected nodes and lines. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey, creating a mesh-like structure.

# Exploratory Data Analysis



# Exploratory Data Analysis

A decorative network graph in the top right corner, featuring a complex web of interconnected nodes and edges. The nodes are represented by circles of varying sizes, some with concentric rings, and the edges are thin lines connecting them. The overall structure is organic and non-linear.

## **Listing features**

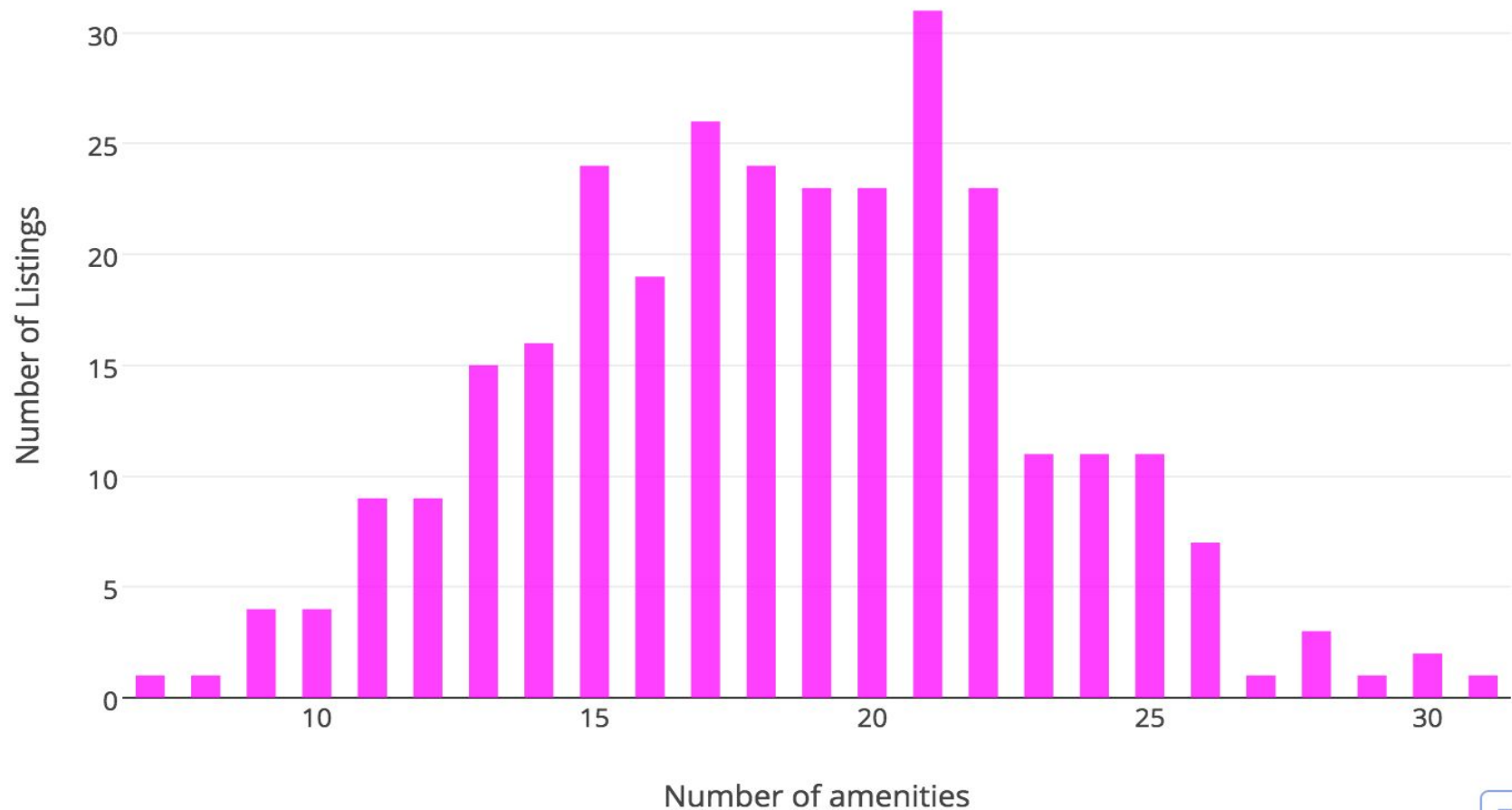
**What is the distribution of amenities over all listings?**

A decorative network graph in the bottom left corner, similar in style to the one in the top right. It shows a cluster of nodes connected by lines, with some nodes having multiple connections, creating a branching or star-like pattern.



# Exploratory Data Analysis

Distribution of listings that offer same number of amenities



[EDIT CHART](#)

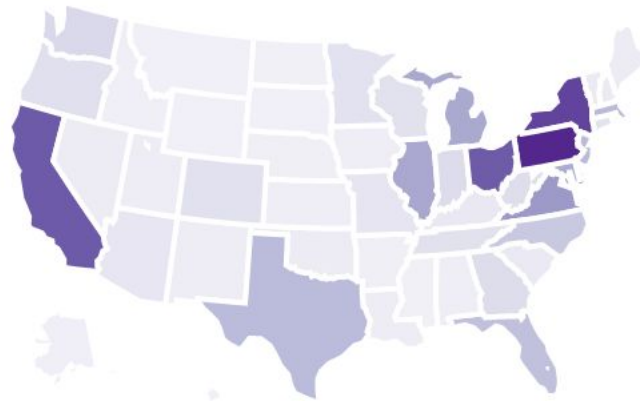
# Exploratory Data Analysis

## **Guests**

**From which part of the world do people visit Pittsburgh?**

# Exploratory Data Analysis

Geography of Airbnb guests in Pittsburgh (within US)



Number of Visitors

1200

1000

800

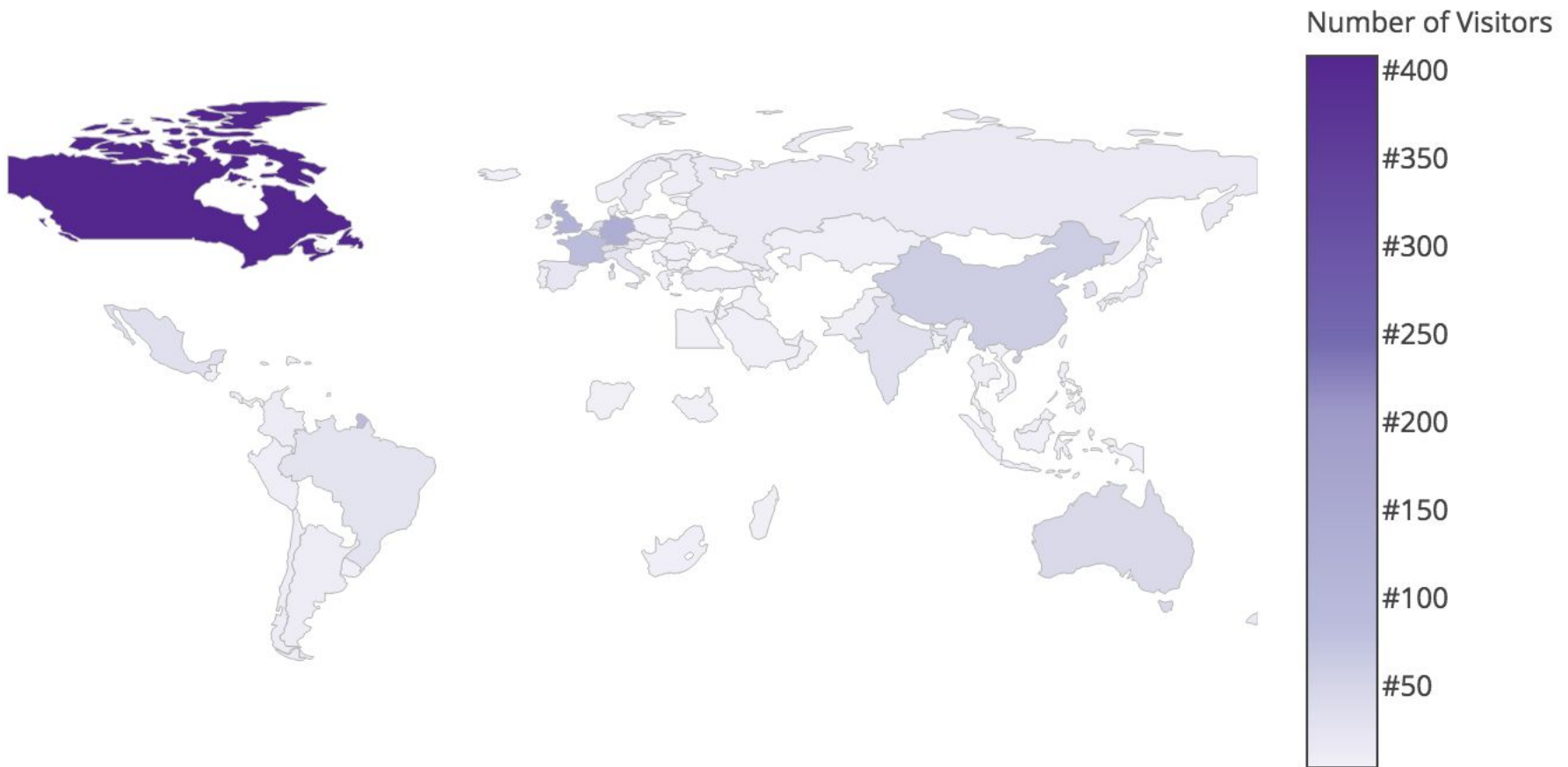
600

400

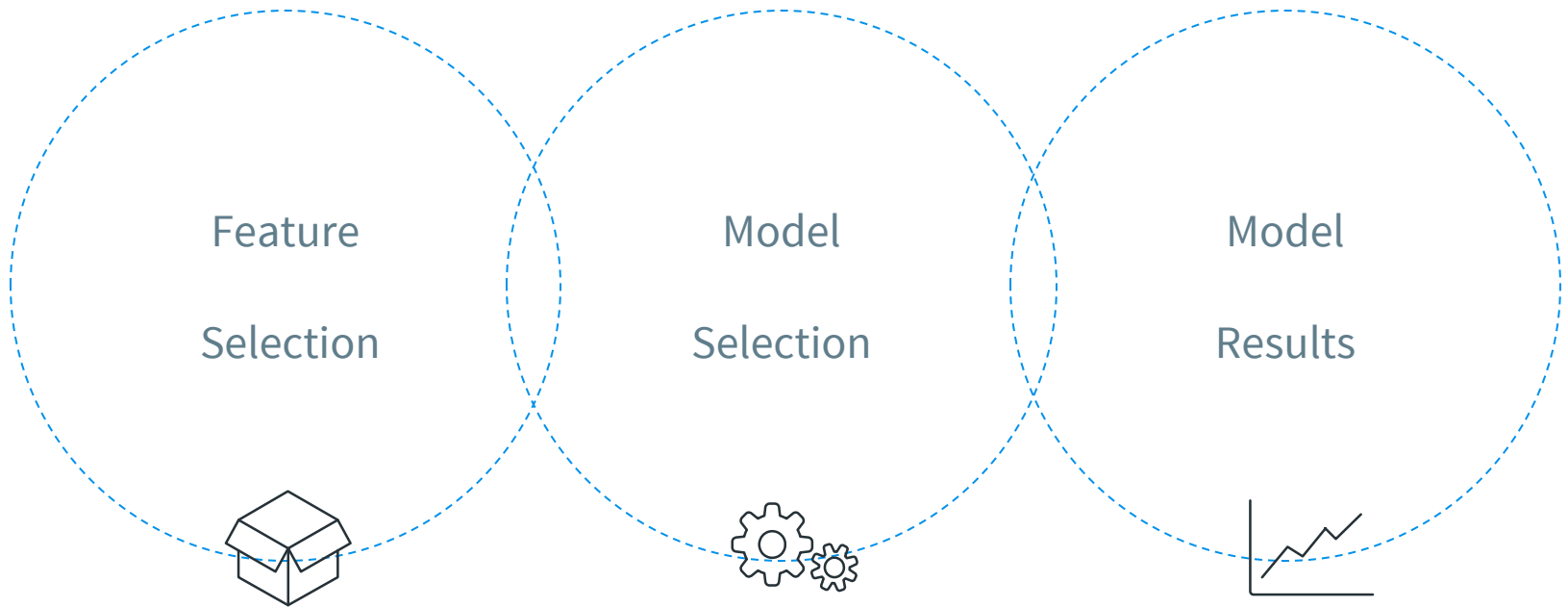
200

# Exploratory Data Analysis

Geography of Airbnb guests in Pittsburgh



# Price Prediction Model



# Feature Selection

## **Listing features**

- lat/lon, #of amenities, price, room type, bedrooms, bathrooms, beds, monthly discount, weekly discount

## **Connectivity & Popularity**

- proximity to different places of interest
- WalkScore API

## **Zone matters**

- Geo clustering

# Feature Selection

## **Listing features**

- lat/lon, #of amenities, price, room type, bedrooms, bathrooms, beds, monthly discount, weekly discount

## **Connectivity & Popularity**

- proximity to different places of interest
- WalkScore API

## **Zone matters**

- Geo clustering
- Add a geo clustering graph here?????

# Model Selection (regression)

## Linear Regression

- Basic (outliers removed)
- Lasso regularization

## Linear Regression with AdaBoosting

- Compensate residuals with a number of linear regressions
- Parameters:

## Gradient boosting regression

- Compensate residuals with some decision trees
- Parameters:



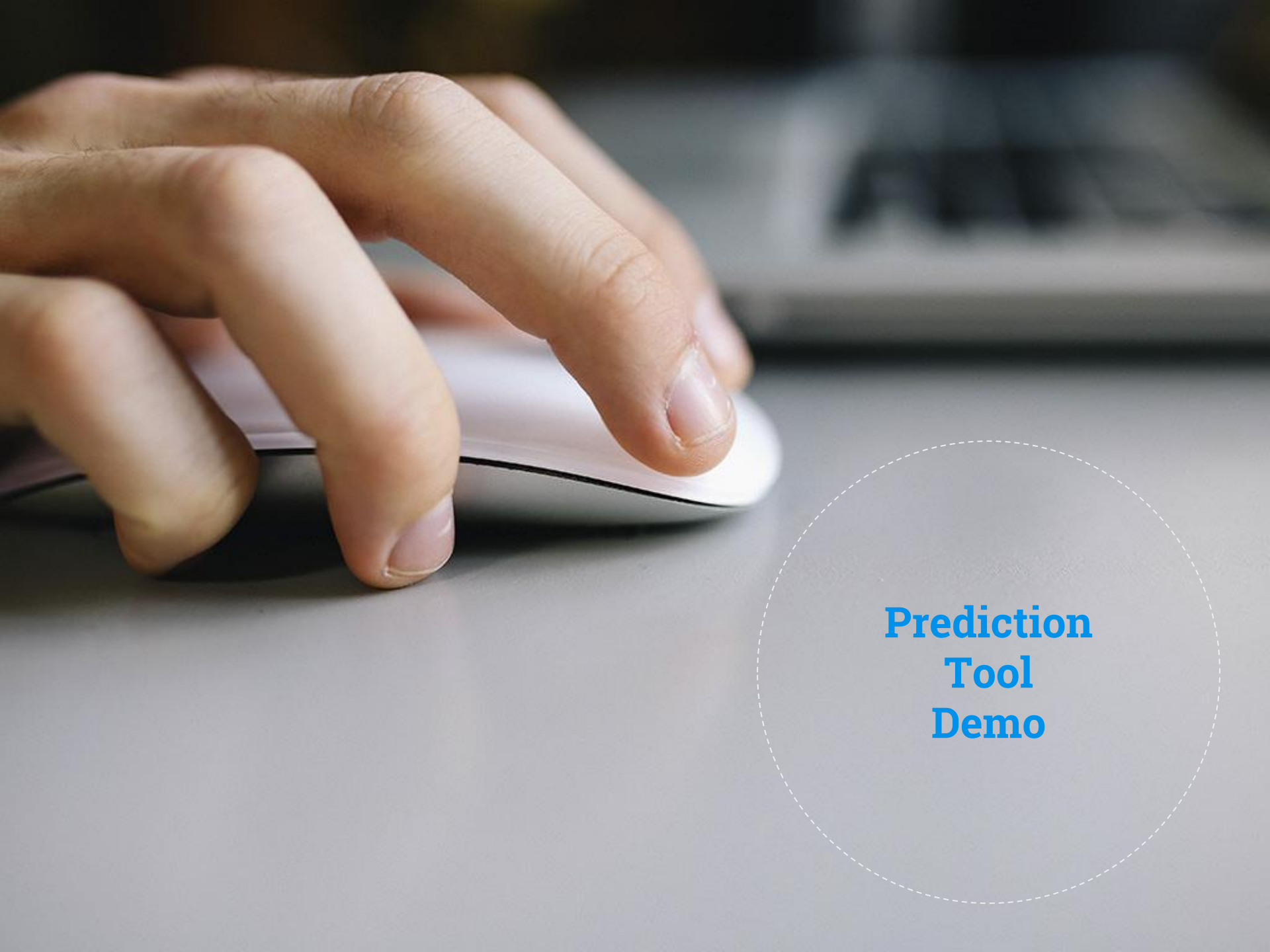
# Model Result

	MSE	R2_Test	R2_Train
Model			
Linear Regression	911.597588	0.334339	0.540044
ADA Boost	543.376868	0.603219	0.720861
Gradient Boosting	766.389933	0.440371	0.984034

Linear Regression with AdaBoost stood out!

## Reasons

- ◎ Price is continuous variable. Using decision tree model to compensate residual is less accurate
- ◎ Max\_depth is small.
- ◎ Scarcity of training data.
- ◎ Sacrifice the benefits of setting more tall trees to avoid overfitting



**Prediction  
Tool  
Demo**

## User Input

- ✓ Listing address
- ✓ Room type
- ✓ Amenities list
- ✓ Bedroom
- ✓ Bathroom
- ✓ Bed
- ✓ Discount



Place your screenshot here

## Prediction Tool Output

- ✓ Predicted Price
- ✓ Pricing stats of nearby Airbnb listings
- ✓ Places of Interest
- ✓ Advertising recommendations



Place your screenshot here

# Location: 8 Mazeroski Way Pittsburgh PA

```
In [40]: #using ada boosting without outliers for our prediction tool
pred = price_prediction_tool(ada,ada_X,kmeans_ada,listings)
```

Address

8 Mazeroski Way Pittsburgh PA

Type:

☐ Entire home/apt

☒ Private room

☐ Shared room

Bedroom

1

Beds

1

Bathroom

1

Amenities

Indoor fireplace

Buzzer/wireless intercom

Suitable for events

Family/kid friendly

Washer

Dryer

Essentials

Safety

Smoke Detector

Carbon monoxide detector

First aid kit

Safety Card

Fire extinguisher

Lock on bedroom door

House Rules

Smoking Allowed

Pets Allowed

24-hour check-in

Self Check-in

Pets

No Pets

Dog(s)

Cat(s)

Other pet(s)

Weekly Discount Factor

0.9

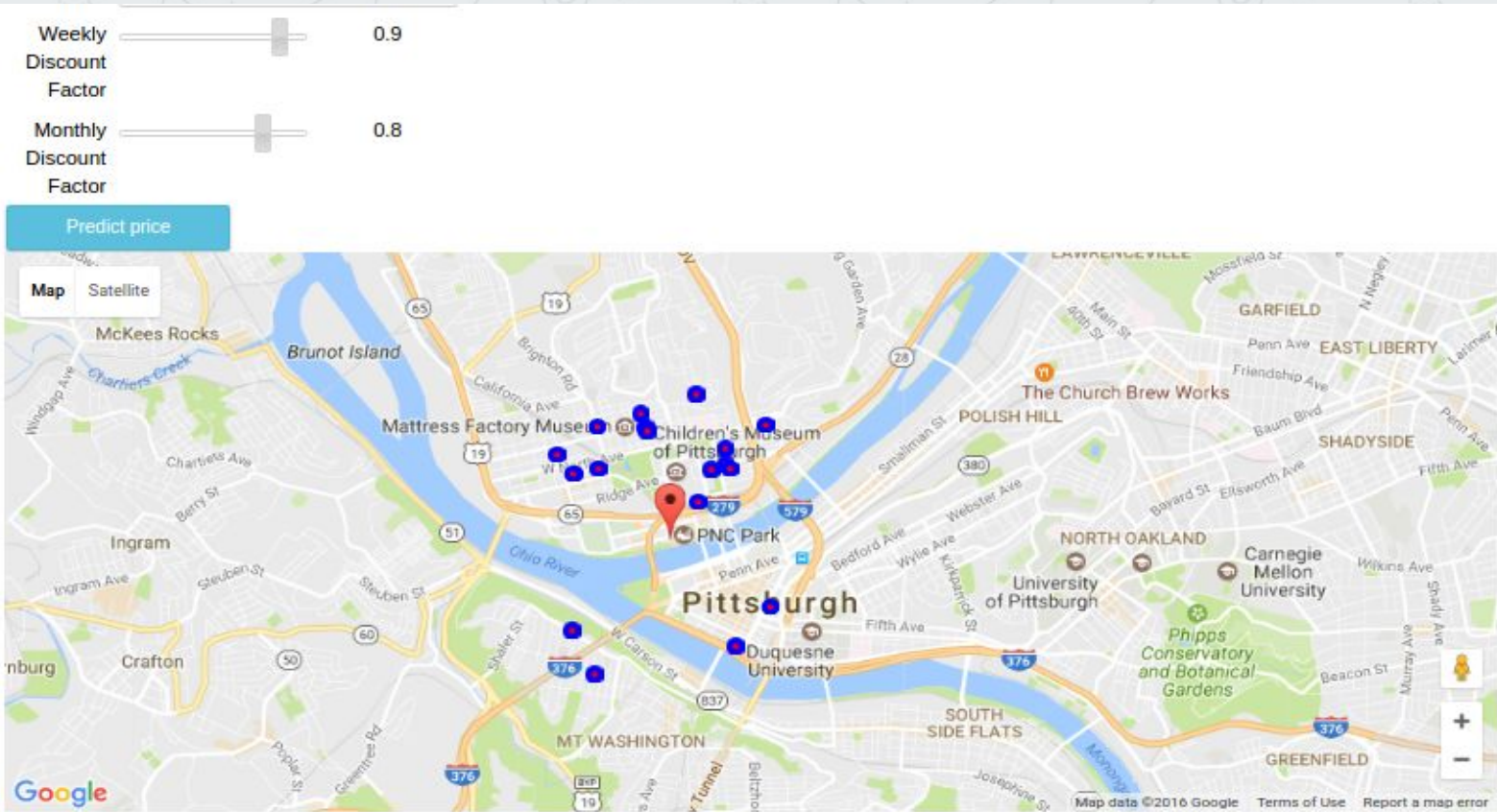
Monthly Discount Factor

0.8

Predict price



# Location: 8 Mazeroski Way Pittsburgh PA



Predicted price for the listing: \$73.91

~~~~~Similar listings nearby~~~~~

Minimum Price: 29.0

Maximum Price: 100.0

Average Price: 57.47

~~~~~Places of Interest~~~~~

Active Life:

North Shore Riverfront Park

Restaurants:

Hyde Park Prime Steakhouse, Slice On Broadway, Lexus Club at PNC Park, Gio's Cafe, Clark Bar & Grill

Entertainment:

PNC Park, Beerhead Bar, Lexus Club at PNC Park, Clark Bar & Grill















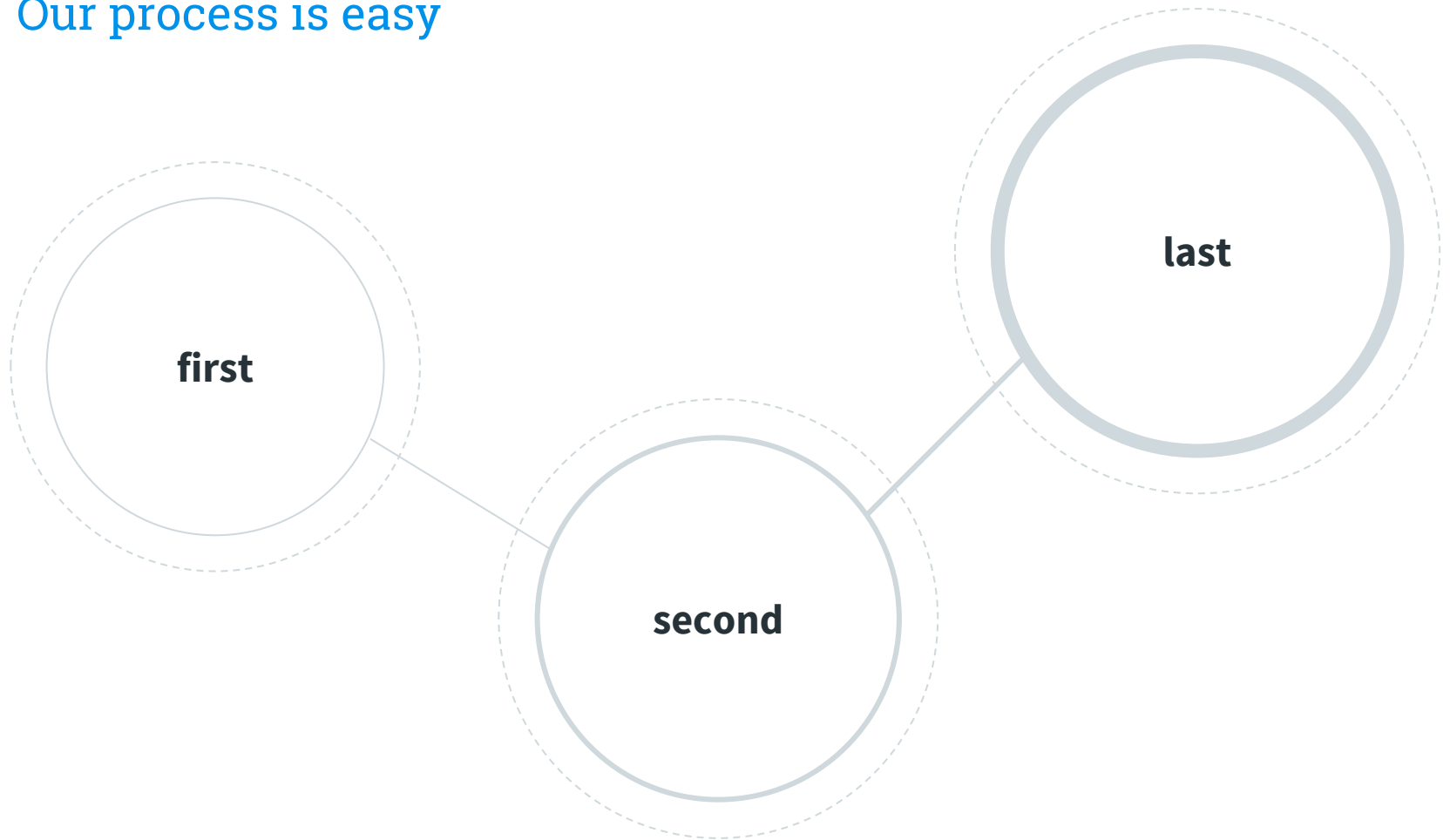


The background of the slide features a complex, light gray network pattern. It consists of numerous small circles, some of which are solid gray and others are hollow with a dashed outline. These circles are interconnected by thin, light gray lines, creating a web-like structure that fills the entire frame. The overall aesthetic is technical and digital.

# 89,526,124

Whoa! That's a big number, aren't you proud?

Our process is easy



## Let's review some concepts



### **Yellow**

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.



### **Blue**

Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.



### **Red**

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.



### **Yellow**

Is the color of gold, butter and ripe lemons. In the spectrum of visible light, yellow is found between green and orange.



### **Blue**

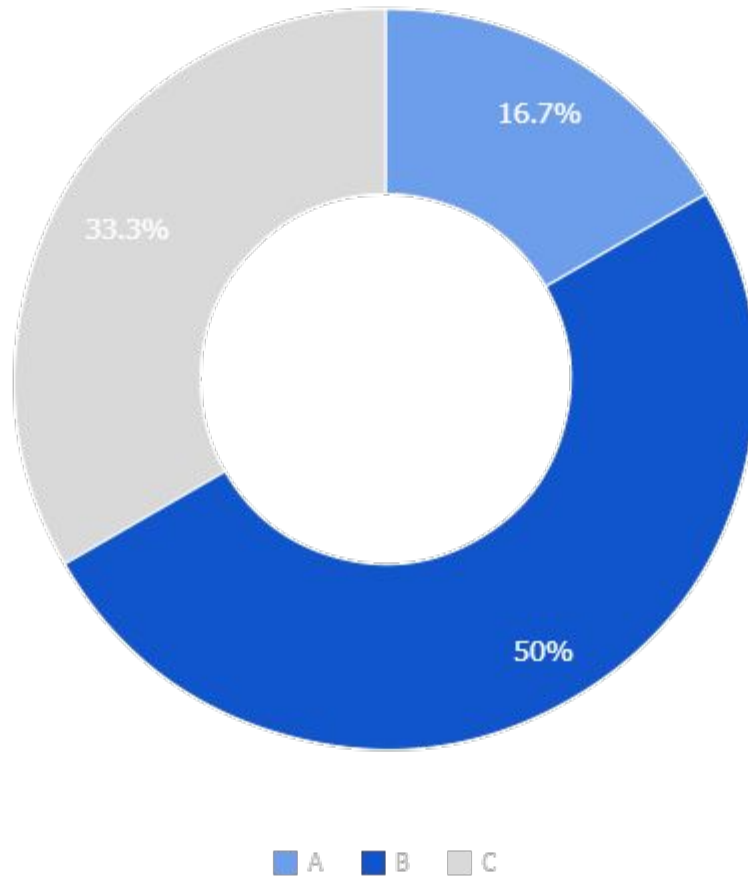
Is the colour of the clear sky and the deep sea. It is located between violet and green on the optical spectrum.



### **Red**

Is the color of blood, and because of this it has historically been associated with sacrifice, danger and courage.



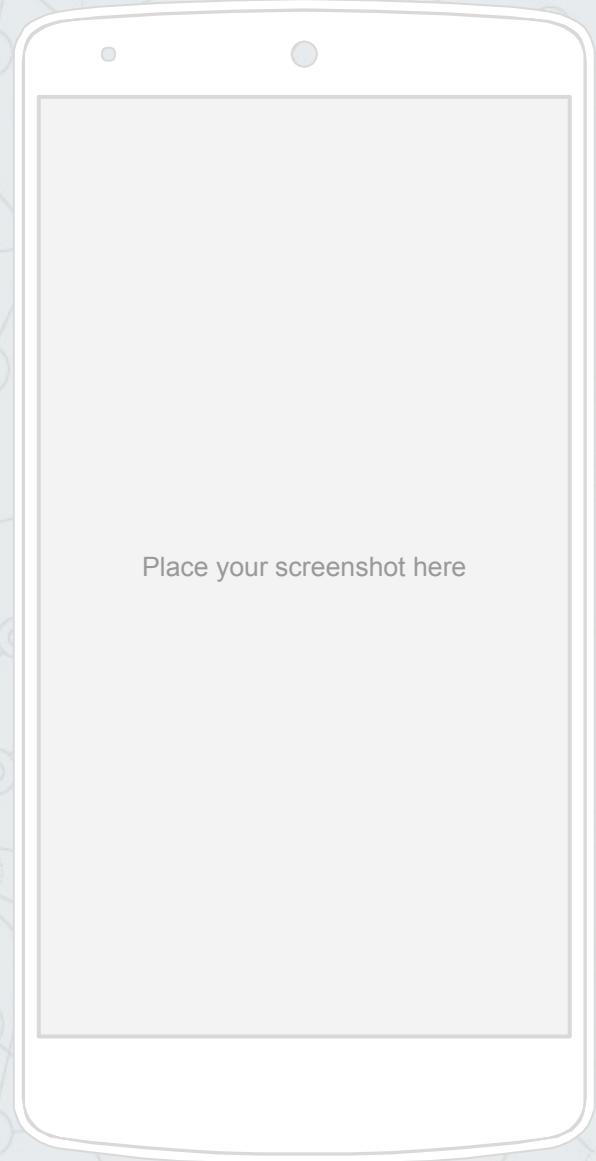


You can copy&paste graphs from [Google Sheets](#)



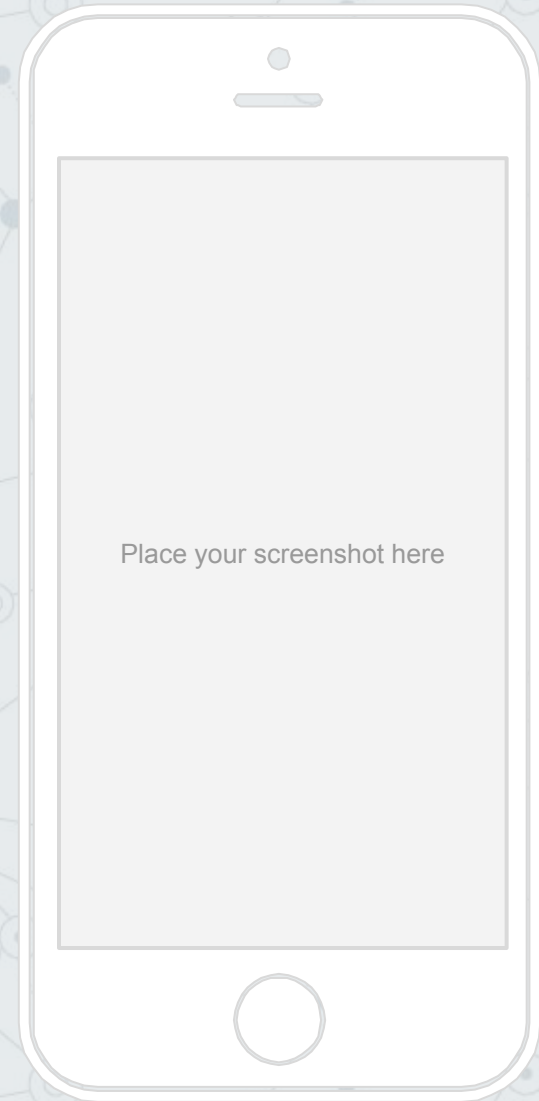
# Android project

Show and explain your web, app or software projects using these gadget templates.



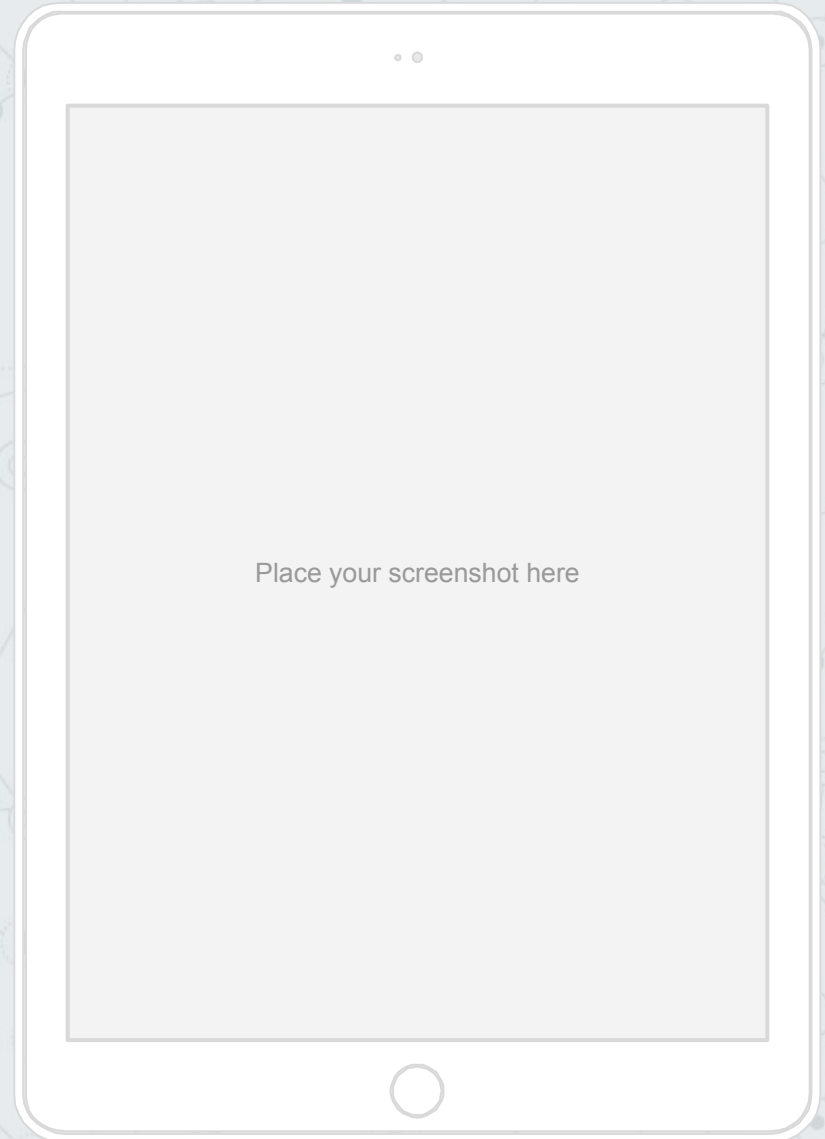
# iPhone project

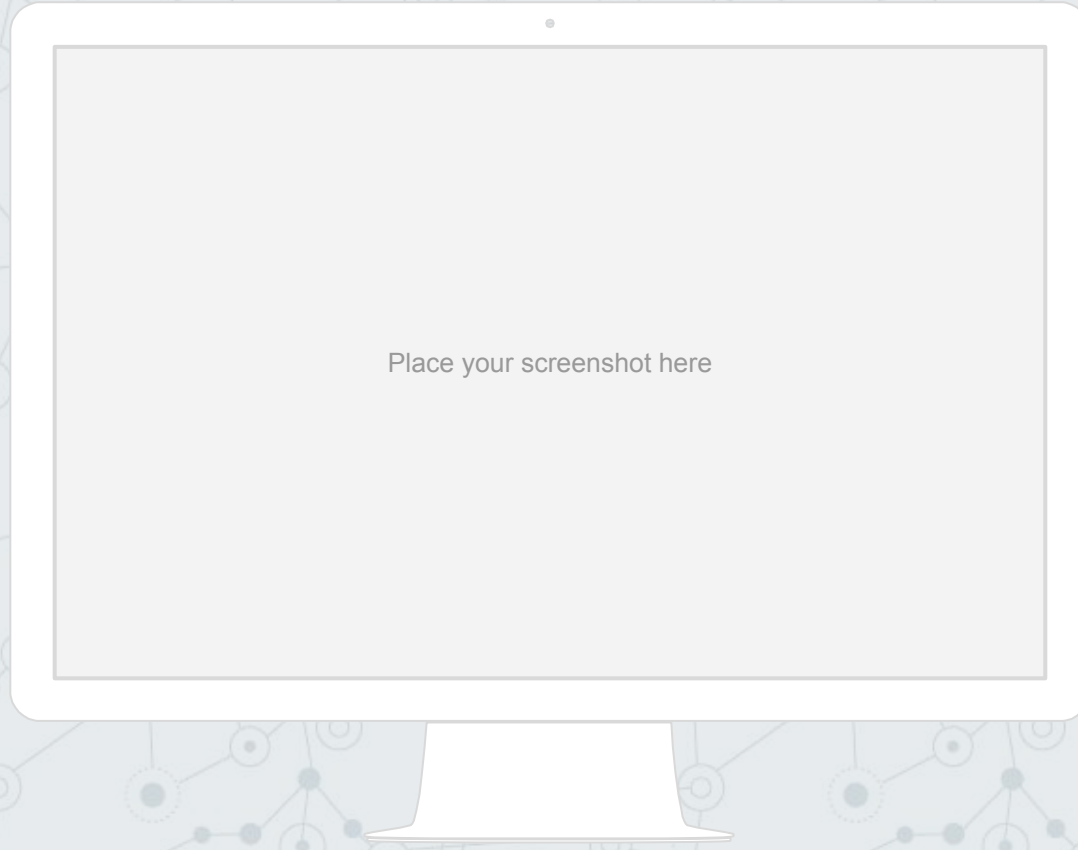
Show and explain your web, app or software projects using these gadget templates.



## Tablet project

Show and explain your web, app or software projects using these gadget templates.





## **Desktop project**


Show and explain your web, app or software projects using these gadget templates.



# Thanks!

## Any questions?

You can find me at:  
@username  
user@mail.me



## Credits

Special thanks to all the people who made and released these awesome resources for free:

- © Presentation template by [SlidesCarnival](#)
- © Photographs by [Unsplash](#) & [Death to the Stock Photo \(license\)](#)

## Presentation design

This presentations uses the following typographies and colors:

- ⦿ Titles: **Roboto Slab**
- ⦿ Body copy: **Source Sans Pro**

You can download the fonts on this page:

<https://www.google.com/fonts#UsePlace:use/Collection:Source+Sans+Pro:400,700,400italic,700italic|Roboto+Slab:400,700>

Click on the “arrow button” that appears on the top right



- ⦿ Blue **#0091ea**
- ⦿ Dark gray **#263238**
- ⦿ Medium gray **#607d8b**
- ⦿ Light gray **#cfd8dc**

*You don't need to keep this slide in your presentation. It's only here to serve you as a design guide if you need to create new slides or download the fonts to edit the presentation in PowerPoint®*





## SlidesCarnival icons are editable shapes.

This means that you can:

- Resize them without losing quality.
- Change line color, width and style.

Isn't that nice? :)

Examples:







## Now you can use any emoji as an icon!

And of course it resizes without losing quality and you can change the color.

How? Follow Google instructions

<https://twitter.com/googledocs/status/730087240156643328>

