Let’s start an Airbnb host career with success!

Advice for Airbnb new hosts

What is Airbnb?

Huge business map

> *The Ebay for the Entire House*
>
> *– Forbes*

Airbnb is based in San Francisco, California, **operates an online marketplace for lodging, primarily homestays for vacation rentals, and tourism activities.** The platform is accessible via website and mobile app.

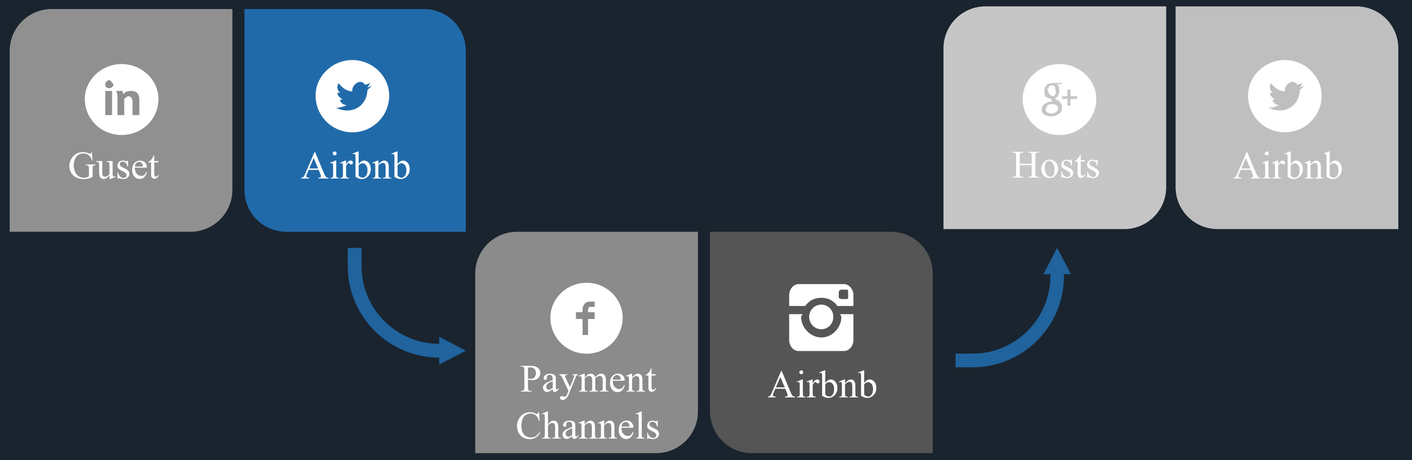
As of September 2020, Airbnb has:

- 5.6 million active Airbnb listings
- in 200+ countries and regions
- 100,000 cities
- 4+ million hosts
- 800+ million guests worldwide.

Business Model

All transactions are done on the Airbnb platform

Success ?
Themselves !



C2C: transactions between hosts (customers) and guests (customers)

Revenue Streams: charges hosts a 3% host service fee per booking, and also adds a 6-12% service fee that’s paid by the guests.

Welcome! Aspiring new Airbnb hosts!

airbnb

optimization

consulting

Too fierce competition?

Price → Success

Longing for a good start?

Walk into CMMQ!

Greatly prove your profit

With

Simple changes in amenities!

Data Understanding

50,600 authentic and reliable records from Airbnb

Geographical Location

Neighborhood they belong to. e.g. Manhattan, Brooklyn, Queens

Room Configuration

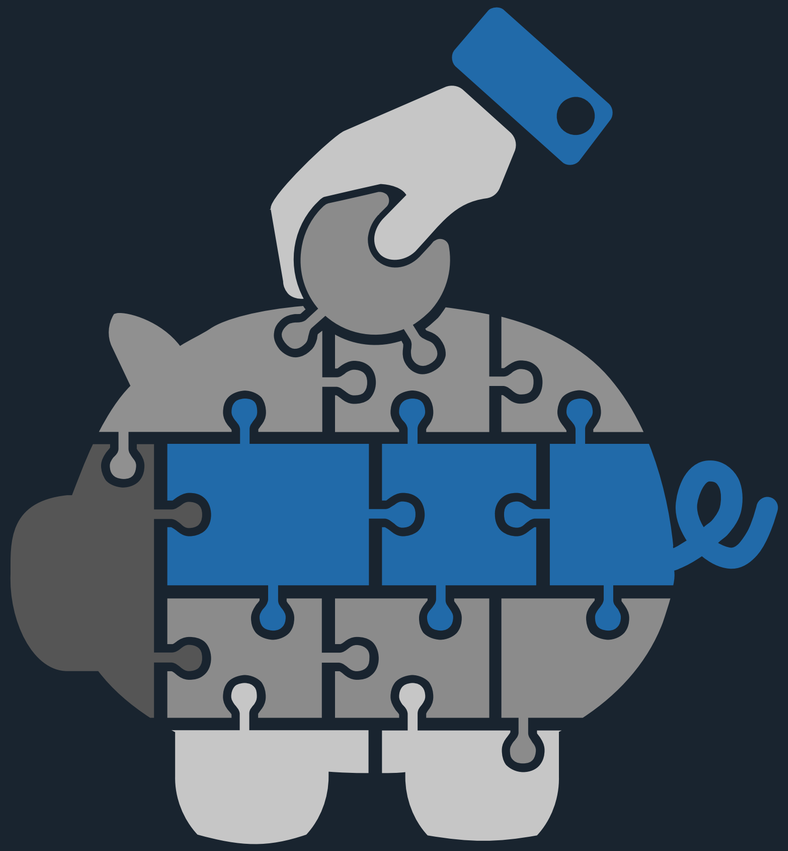
including house structure, amenities, etc.

Hosts Information

e.g. calculated_host_listings_count, host_response_rate, host_response_time_norm, etc

Lease information

including price, number of days, etc.



Data Preparation

How we get useful data

Feature Selection (part I)

Select features related to the model

Data Processing (part I)

- Missing value filling
- Digitize and normalize strings
- Creating sparse matrixes for some features

Feature Selection (part II)

For the amenity sparse matrix:

- Combine similar and duplicate features
- Delete features that only have a single value (both are 0)

Data Processing (part II)

price has extreme values

Separate the training set and test set

according to mean + variance values

filter out examples whose price falls outside the normal range.

Data set division

according to whether

calculated_host_listings_count>=1

- Training Set (the data of experienced/old hosts)
- Test Set (the data of new hosts)

Output training set, test set

Modeling

Target Value: Price Tool: Python

Algorithms:
linear regression, logistic regression, SVR, ridge regression, Elastic Net, K-Neighbors Regression, Random Forest Regression, Neutral Network, Gradient Boosting Regression

Evaluation:
mean_absolute_error (MAE), mean_squared_error (MSE), median_absolute_error, r2_score, Coefficient of Determination

Step 1: New Data Set
Old host data → new training set + new test set
ratio: 80%

Step 2: modeling and Evaluation

Step 3: Evaluation
Evaluate the new test set (the data of old hosts)
Verify the new host test set

Step 4: Comparison of results

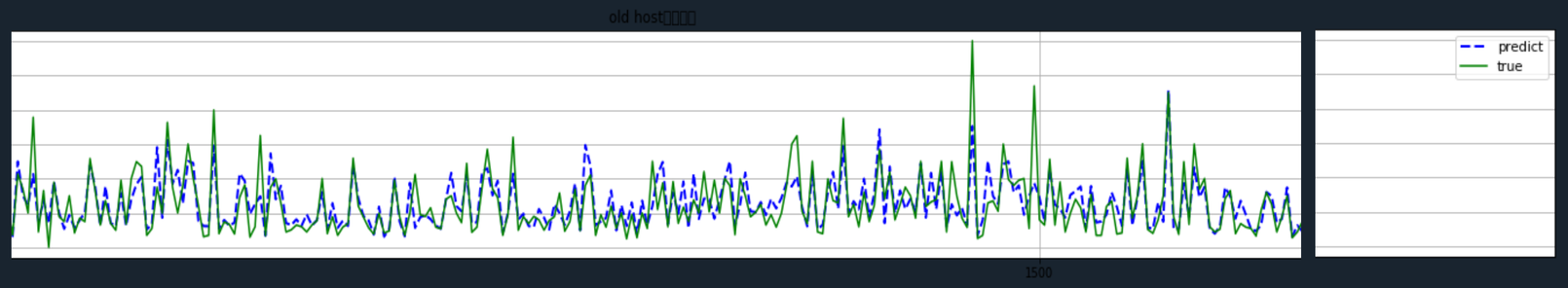
Comparison of Results

		MAE	MSE	median_absolute_error	r2_score	Coefficient of Determination
linear regression						
	old host test set	44.01	4784.24	27.98	0.49	0.49
	new host test set	48.60	4655.40	35.13	0.38	0.38
logistic regression						
	old host test set	45.13	6097.42	23.00	0.39	0.35
	new host test set	63.53	8750.13	40.00	-0.16	-0.17
Random Forest Regression						
	old host test set	34.49	3403.76	18.50	0.64	0.64
	new host test set	53.24	5364.81	37.70	0.31	0.28
K-Neighbors Regression						
	old host test set	38.93	4437.12	20.80	0.53	0.53
	new host test set	56.53	6385.90	38.60	0.15	0.15
SVR						
SVR（outliers 0.01 are not removed）						
	old host test set	42.70	5370.46	23.73	0.46	0.43
	new host test set	48.48	5285.56	30.40	0.36	0.29
SVR（outliers 0.06 are not removed）						
	old host test set	36.78	4267.71	18.95	0.57	0.55
	new host test set	48.90	5026.71	32.84	0.36	0.33
SVR（outliers 0.03 are not removed）						
	old host test set	39.73	4867.72	20.49	0.51	0.48
	new host test set	48.00	5103.76	30.56	0.38	0.32
Neutral Network						
	old host test set	42.04	4367.35	26.08	0.54	0.54
	new host test set	50.55	4757.25	37.59	0.37	0.36
ridge regression						
	old host test set	43.92	4791.91	27.70	0.49	0.49
	new host test set	48.46	4655.94	34.96	0.38	0.38
Elastic Net						
	old host test set	45.02	5046.27	28.68	0.47	0.47
	new host test set	49.92	4909.89	35.73	0.35	0.34
Gradient Boosting Regression						
	old host test set	39.29	3991.34	24.48	0.58	0.58
	new host test set	46.92	4295.40	33.61	0.43	0.43

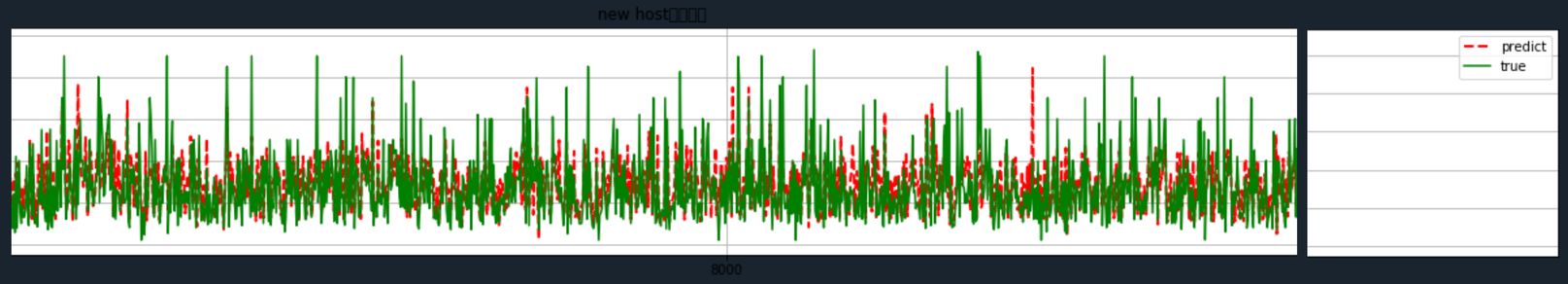
Random Forest Regression

Random Forest Regression		MAE	MSE	median_absolute_error	r2_score	Coefficient of determination
	old test set	34.49	3403.76	18.50	0.64	0.64
	new test set	53.24	5364.81	37.70	0.31	0.28

Old host test set:



New host test set:



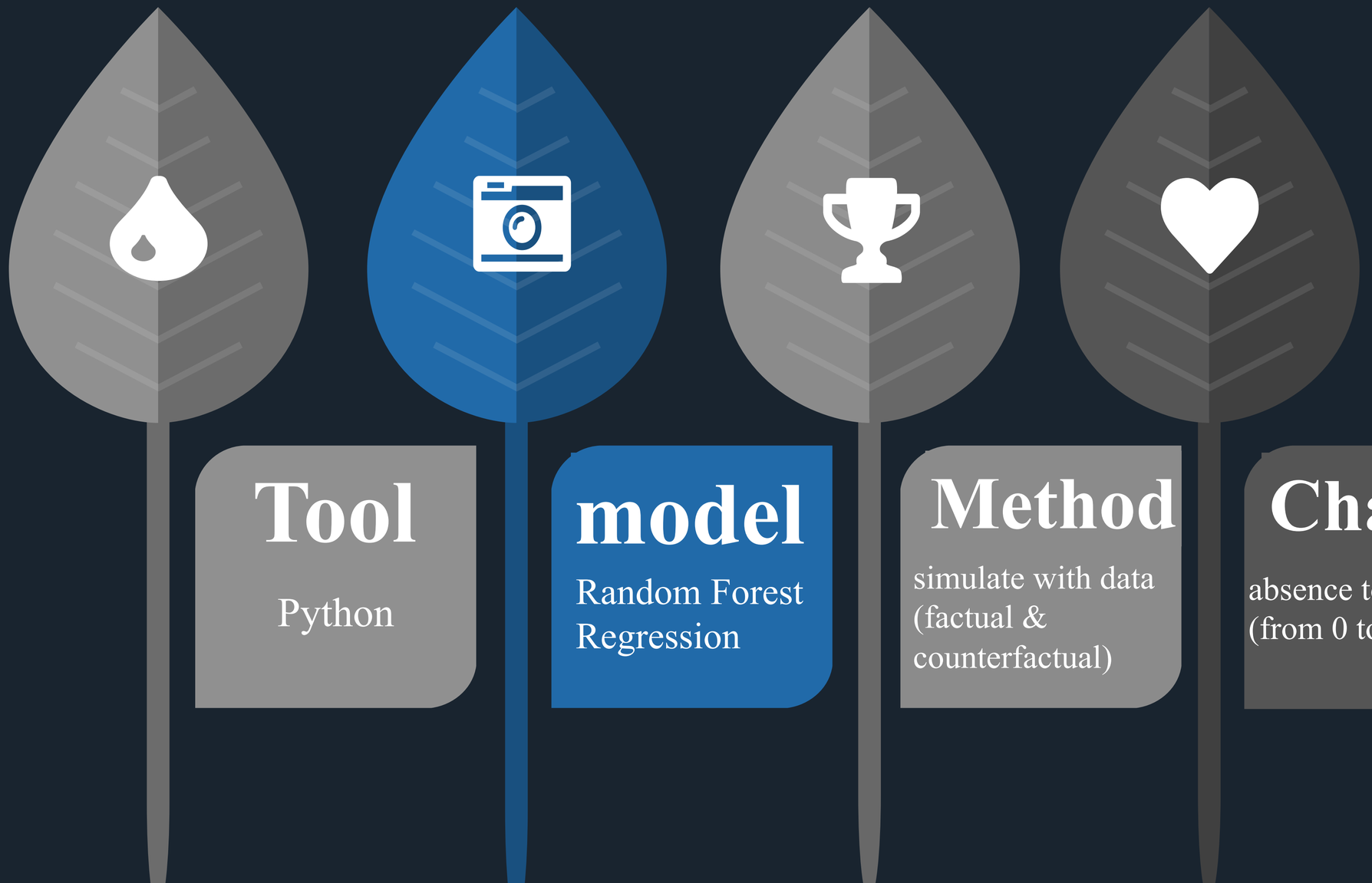
Ranking

More amenities, higher prices

	Amenities	Ratio		Amenities	Ratio		Amenities	Ratio		Amenities	Ratio
1	Suitable for events	10.22%	8	Beach essentials	4.12%	15	bathtub	1.82%	21	Hot water kettle	0.75%
2	baby care	8.18%	9	Private living room	4.05%	16	TV	1.49%	22	Cleaning before checkout	0.56%
3	Dishwasher	7.77%	10	Shampoo	3.63%	17	Luggage dropoff allowed	1.44%	23	disabled friendly facilities	0.51%
4	Smoking allowed	7.65%	11	pets allowed	3.43%	18	Air conditioning	1.39%	24	Oven	0.51%
5	Breakfast	7.32%	12	family/kids-friendly	3.14%	19	Hangers	0.83%	25	Comfortable sleeping environment	0.44%
6	washer/dryer	5.85%	13	Hair dryer	2.62%	20	Game console	0.79%	26	First aid kit	0.38%
7	Indoor fireplace	5.83%	14	BBQ grill	2.61%		Long term stays allowed	0.78%	27	Iron	0.34%
Ratio=average [(price with amenities / price without amenities) - 1]									28	workspace	0.30%

Is that useful?

The impact of amenity on price



Tool

Python

model

Random Forest Regression

Method

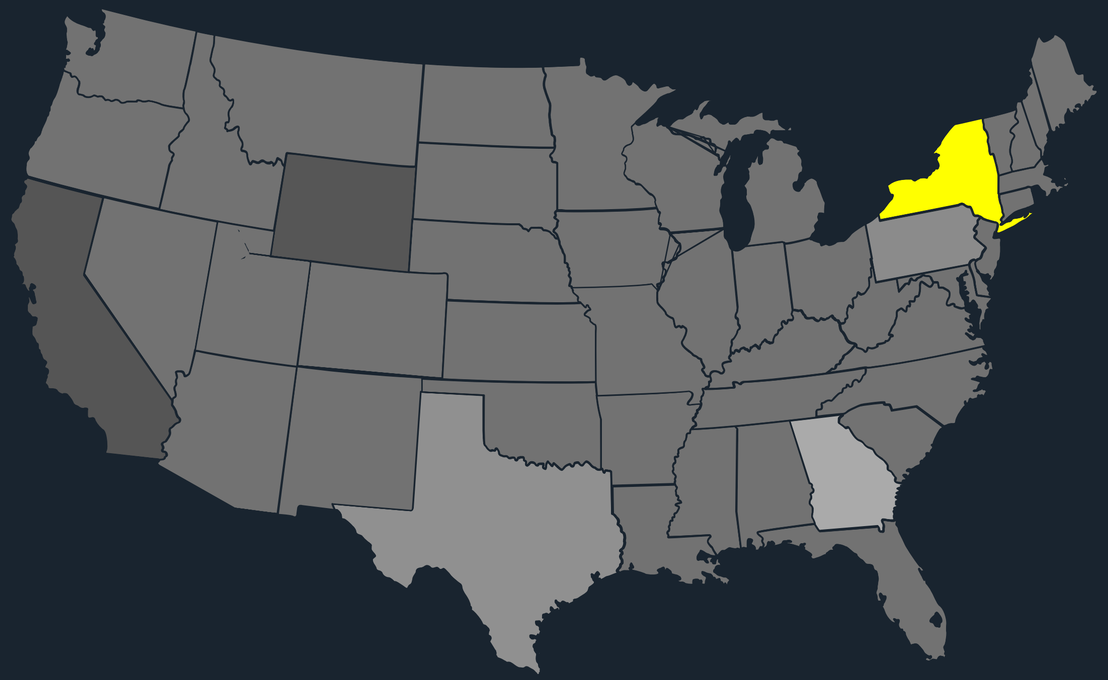
simulate with data (factual & counterfactual)

Change

absence to presence (from 0 to 1)

Use Case

Let us help a new host get a higher price!



Max
Queens
pregnant —> make extra money
have an empty private room
Amenities:
The fire and gases prevention & TV

Price: $39

Change Time!

More profit

Before Change

Predicted Price: $57.5

Cost/yr: $0
Predict Revenue: $20987.5
Predict Net Profit: $20987.5

Optimization Plan —— Add Amenities

Long term stays allowed, Heating, Workspace, Private living room, Indoor fireplace, Refrigerator, Breakfast, Beach essentials, Coffee maker, Hair dryer, Shampoo, Luggage drop off allowed, Dishwasher, Smoking allowed, Cleaning before checkout, Game console, Iron, First aid kit, Dishes and silverware, Hot water kettle, Hangers, Air conditioning, BBQ grill, Suitable for events, Cozy bathroom, Internet, Comfortable sleeping environment, good check-in experience, disabled friendly facilities, bathtub, pets allowed, cooking basics, family/kids-friendly, baby care, washer/dryer, Essentials, Oven

After Change

Predicted Price: $171.8

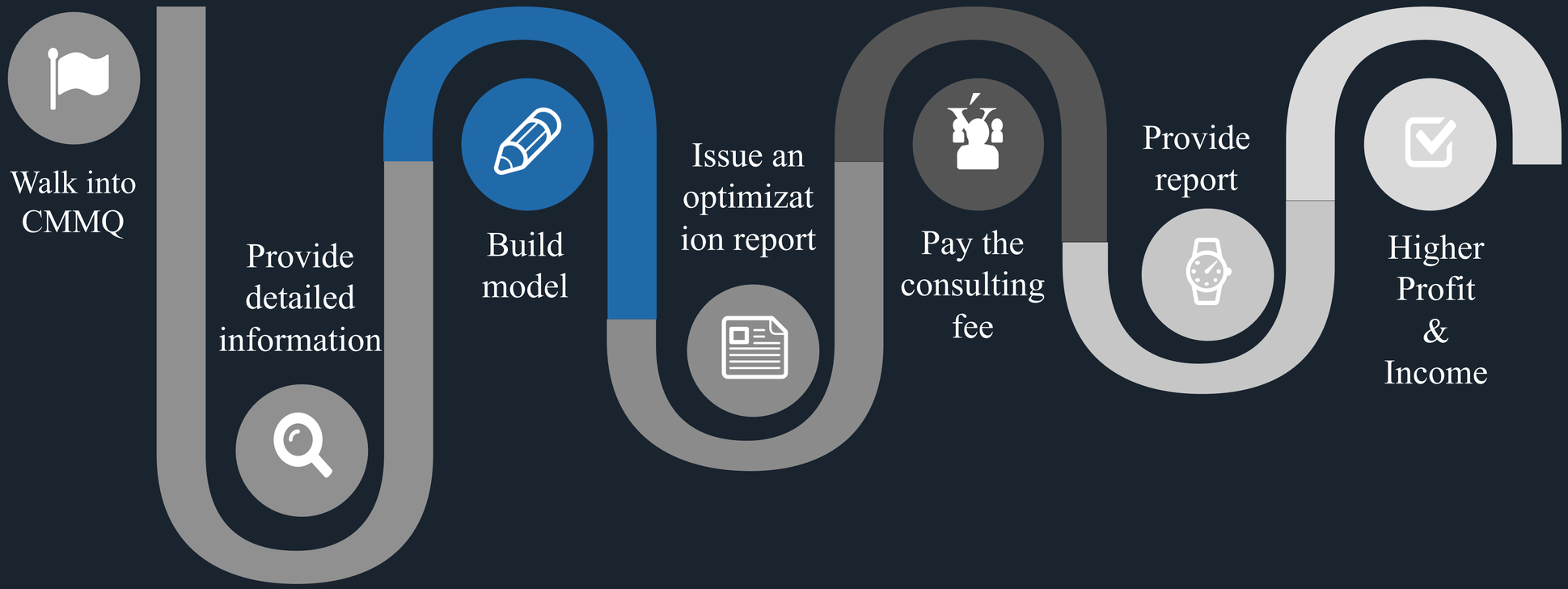
Cost/yr: $631.6943095
Predict Revenue: $62707
Net Profit: $62075.3057

Net Profit Premium: $41087.8057

Annual Premium: 41719.5

How we Profit

Consultation Fee



Consulting Fee = Net Profit Premium* 20%

Win - Win

Higher Revenue & Income

For Customers

Annual Profit premium (Pre-tax)
= Net Profit Premium * 0.8
= $32870.24

For our company

Profit for this order (Pre-tax)
= Net Profit Premium * 0.2 – 2 Day Salary
= $5041.26

*Salary/day =3*$377.296 (Consultant) + $456.248 (Data Scientist)

Limitations

Some shortcomings of our model

- Different customer groups have different needs for different amenity, but we lack of customer segmentation in the data set.
- The hosts' own classification of amenities is not rigorous enough different hosts have different understandings of different amenities.

Thanks for listening

Appendix - Business Canvas

<table>
<tr><th>key partners</th><th>key activities</th><th>Value proposition</th><th>Customer Relationship</th><th>Customer segments</th></tr>
<tr>
<td rowspan="3">1. Hosts
2. Guests
3. Real Estate Developers
4. Local Governments
5. Industrial Organization
6. Payment Channels
7. Technical Experts
8. Insurance
9. Marketing
10. Service companies</td>
<td>1. marketing
2. Update and maintain the website and APP
3. Attract potential users
4. Product Segmentation
5. User Service
6. Community Operation
7. Product Development</td>
<td rowspan="3">1. For Guests
1) convince
2) lower price
3) Opportunities to integrate into the local community
4) Variety of choices
5) Unique travel experience
2. For Hosts
1) Extra income
2) convince
3) Risk aversion</td>
<td>1. Marketing
2. Effective problem solving
3. Build an Airbnb community
4. Superhost program
5. Safety
6. User-friendly Product</td>
<td rowspan="3">1. Hosts
1) part-time Airbnb hosts
2) full-time Airbnb hosts
3) real estate developers
4) property managers
2. Guests
1) Price-sensitive guests
2) Environmentally sensitive guests
3) Experience-sensitive guests
4) Amenity sensitive guests
5) Service-sensitive guests</td>
</tr>
<tr>
<td>Key Resources</td>
<td>Channels</td>
</tr>
<tr>
<td>1. Brand
2. Hosts
3. Guests
4. Employees
5. Network Platform
6. Algorithm
7. Data Sets</td>
<td>1. Website
2. Mobile APP
3. Social Media
4. Offline Advertisement
5. Online Advertisement
6. Word-of-mouth
7. promotional discounts</td>
</tr>
<tr><th colspan="3">Cost Structure</th><th colspan="2">Revenue streams</th></tr>
<tr><td colspan="3">1. Transaction fees for online payment 2. Maintenance cost of website and APP 3. R&D cost 4. Insurance 5. Marketing 6. Labor costs 7. Cost of sales</td><td colspan="2">Airbnb charges hosts a 3% host service fee per booking, and also adds a 6-12% service fee that’s paid by the guests.</td></tr>
</table>