

“Airbnb Pricing: Identifying Primary Determinants through Hedonic Pricing Methods”

Konstantinos Gkovedaros

Digital Analytics Presentation

Wednesday, 22nd of May, 2019



Who am I?

- Engineer with an understanding of economics
- Not an economist
- Not a data scientist
- Basic use of R
- *Tools used: R, Excel, Google Cloud – Distance Matrix API, Eviews,



Introduction

What is Airbnb?

- World's largest accommodation-sharing marketplace
- 5 million listings, 81.000 cities, 191 countries
- 150m users, 640k hosts
- Founded in 2008, current valuation estimated at \$31bn
- Approx. \$4bn yearly revenue (2018)



Introduction

What started

Τα νέα των μικρομεσαίων επιχειρήσεων

Ταχύτατα αναπτυσσόμενη στην Ελλάδα η βραχυχρόνια μίσθωση ακινήτων τύπου Airbnb

Airbnb: Η Νέα Πέραμος Καβάλας στους πιο ανερχόμενους | Απογοητευμένη η πλειοψηφία των εμπόρων από την

Δημοσίευση: 07/05/2019 - 17:55 | Τελευταία ενημέρωση: 07/05/2019 - 17:55

← | Σχολιάστε

και χαοτικό το νομικό «φαινόμενο Airbnb»: Καθηγητές προβλήματα

...ηθούν το «φαινόμενο Airbnb» και πώς

ΟΙΚΟΝΟΜΙΑ 10/05/2019 15:12 EEST | Updated 10/05/2019 15:13 EEST

Πάνω από 340.000 τουρίστες μέσω Airbnb στην Αθήνα το 2018 - Εσοδα 90 εκατ. ευρώ

Μόλις το 5,3% των ενοικιαζομένων κατοικιών στην Αττική δηλώθηκαν στην πλατφόρμα βραχυχρόνιας μίσθωσης.

Newsroom
HuffPost Greece

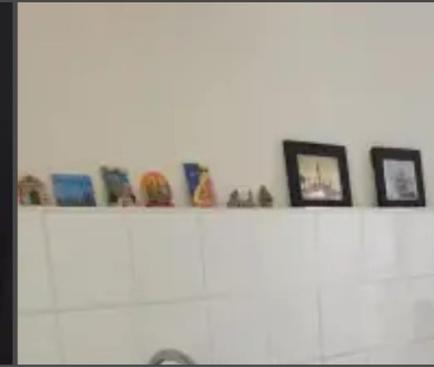
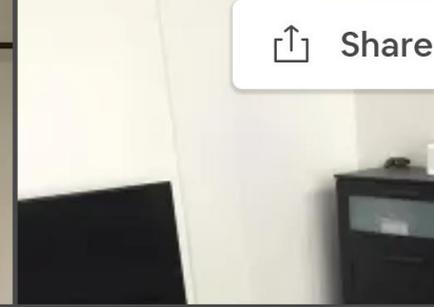


So the question is this

*What do guests value most
in Airbnb accommodations?*



Introduction



Chambre privée quartier européen

Bruxelles

🏠 Private room in apartment

2 guests 1 bedroom 1 bed 1 shared bath

📍 Great check-in experience

95% of recent guests gave the check-in process a 5-star rating.

🌟 Véronique is a Superhost



Véronique

Add dates for prices

★★★★★ 139

Dates

Check-in



Checkout

Guests

1 guest



Add dates

Les chambres sont à l'étage.

La salle de bain et wc sont communs et se situent au rez de chaussée.

[Hide](#) ^

Contact host

Amenities



Wifi



Laptop friendly workspace



Iron



Hangers



~~Carbon monoxide detector~~

The host hasn't reported a carbon monoxide detector on the property.

[Show all 20 amenities](#)

Sleeping arrangements

Add dates for price

★★★★★ 139

Dates

Check-in

Guests

1 guest

[Add](#)

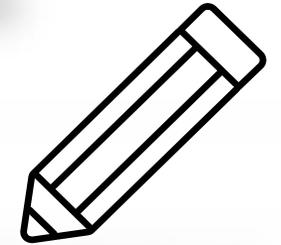
This home is on people's

It's been viewed 141 time
week.

Report

Pricing in the hospitality industry

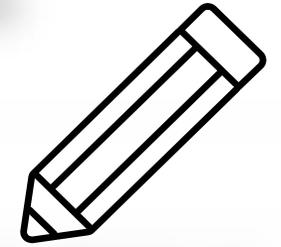
- pricing and revenue management are two of the most frequently researched subjects in hospitality marketing, but not for Airbnb (*Yoo, Lee, & Bai, 2011*)
- the growth in the use of the Internet for travel planning and purchasing has made pricing more transparent (*Gibbs et al., 2017*)
- though price is important, Airbnb hosts are confounded when setting one (*Hill, 2015*)



Literature Review (1/3)

Airbnb and pricing: What matters (1/2)

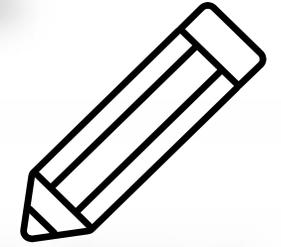
- a number of independent variables related to trust is affecting Airbnb prices considered (*Ert, Fleischer, and Magen, 2016*)
- hosts awarded a “Superhost” badge, post their properties at higher prices, especially when they receive more reviews and higher ratings (*Liang, Schuckert, Law & Chen 2017*)
- common traditional accommodation attributes like location and size clearly matter greatly, as well as privacy (*Gibbs et al., 2017*)



Literature
Review (2/2)

Hedonic pricing models

- “...The simple method is inapplicable, but why not combine several specifications to form a single composite measure?...” (*Court, 1939*)
- The hedonic price model posits that goods are typically sold as a package of inherent attributes (*Rosen 1974*)
- the hedonic method is an "indirect" valuation method in which we do not observe the value consumers have for the characteristic directly, but infer it from observable market transactions (*Champ, Boyle and Brown, 2003*)



Literature
Review (3/3)

Why Athens?

- available dataset from insideairbnb.com
- Athens has a more familiar profile, the legal framework is well-known as well as regulation in place
- there is a considerable buzz in the last few years about Airbnb in Athens and policy-making is eminent
- Athens has been included in empirical research of Airbnb pricing (Wang and Nicolau, 2017) but not examined as a stand-alone case study

Data

- **9.123 listings** with **96 variables** from *insideairbnb.com* scraped on **November 15, 2018**
- among those: listing size, location, amenities, host attributes, reviews score and comments, availability, cancelation policy, etc.
- Only Kentrikos tomeas Athinon taken into account
- remove listings with blank values on price, no reviews, and pricing over €200, left with **6.885 listings**

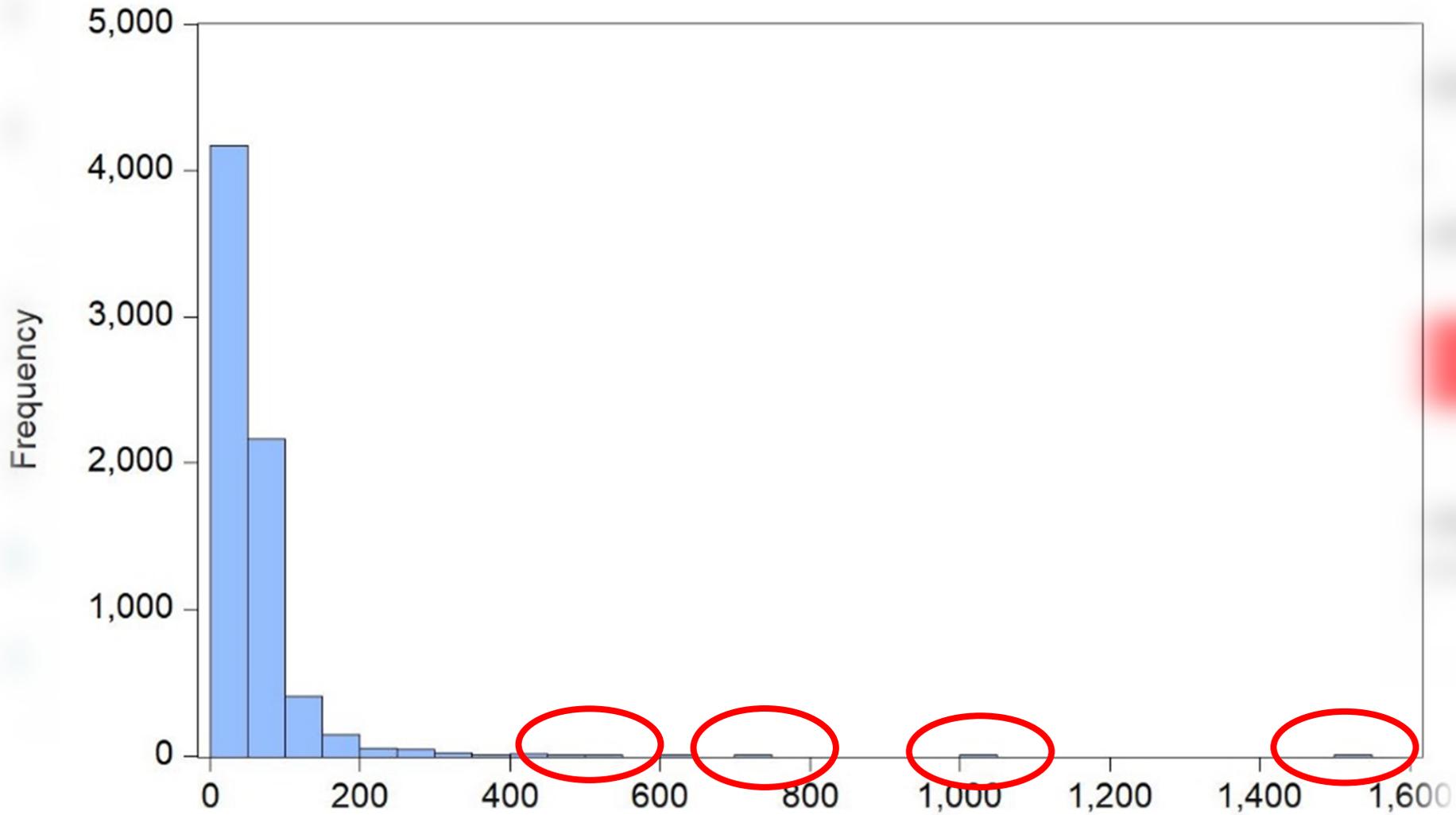


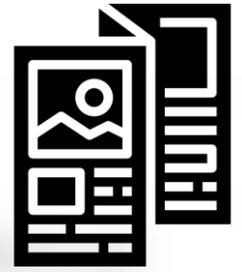
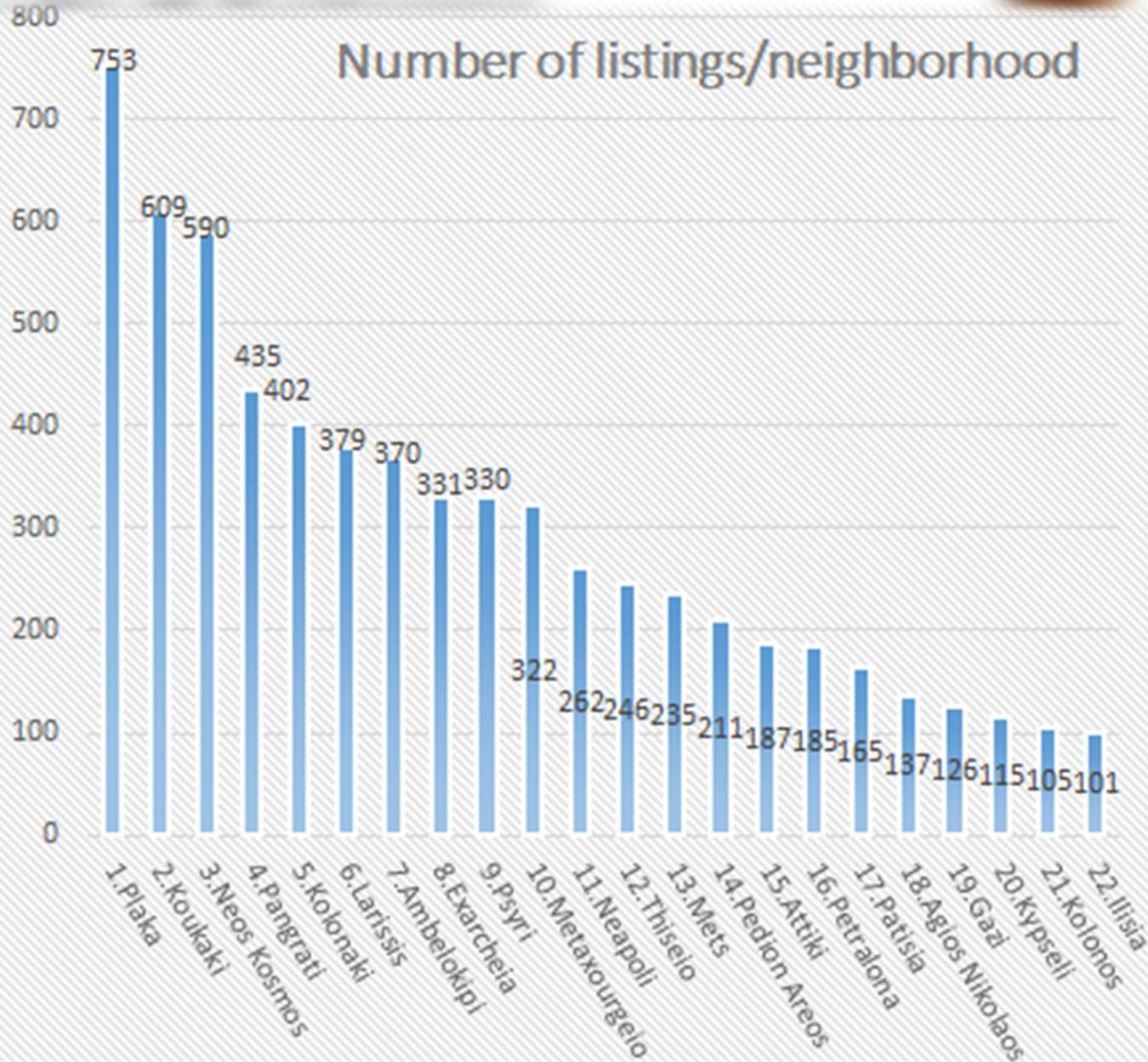
Data



Data

price





Descriptive statistics

About distance (1/3)

Google Cloud Platform Select a project



Distance Matrix API

Google

Travel time and distance for multiple destinations.

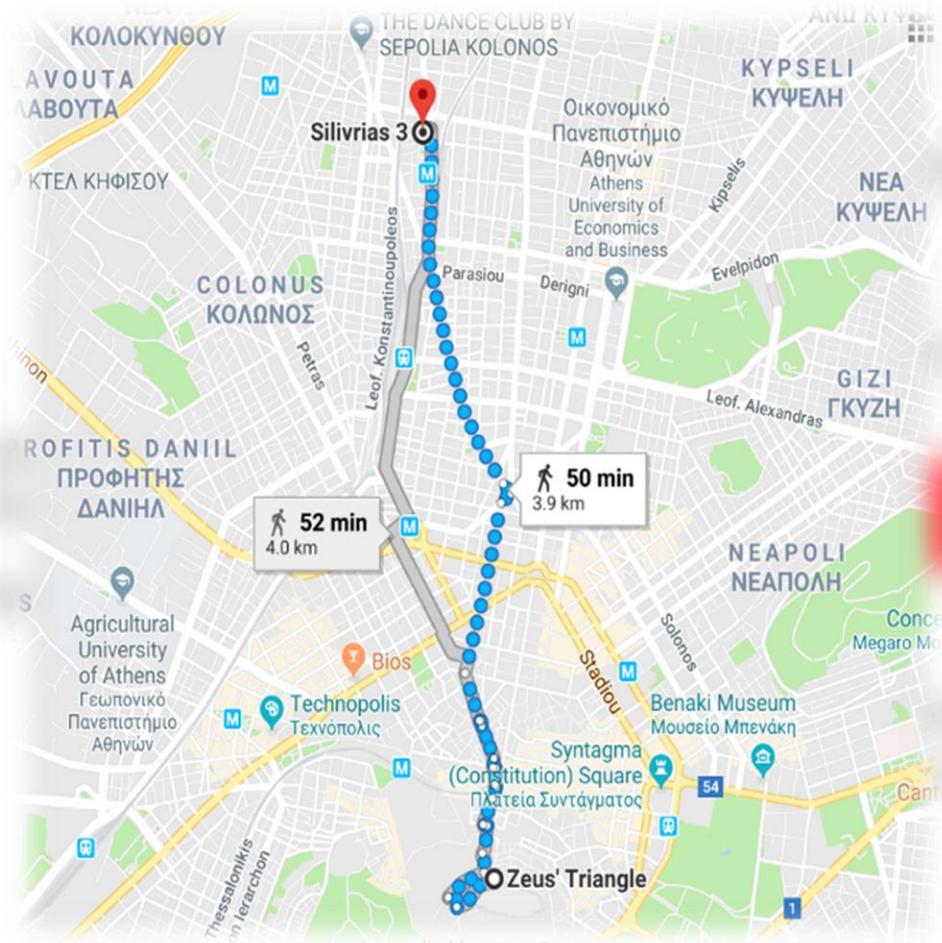
ENABLE



Descriptive
statistics:
Distance

About distance (2/3)

Navigation app interface showing route options between Zeus' Triangle, Athina 105 58 and Silivrias 3, Athina 104 45. The interface includes a menu, transport mode icons (car, train, walking, bicycle, airplane), and a list of destinations. A blue bar at the bottom indicates the selected route via Liosion.



Descriptive statistics:
Distance

via Liosion 50 min
DETAILS 3.9 km

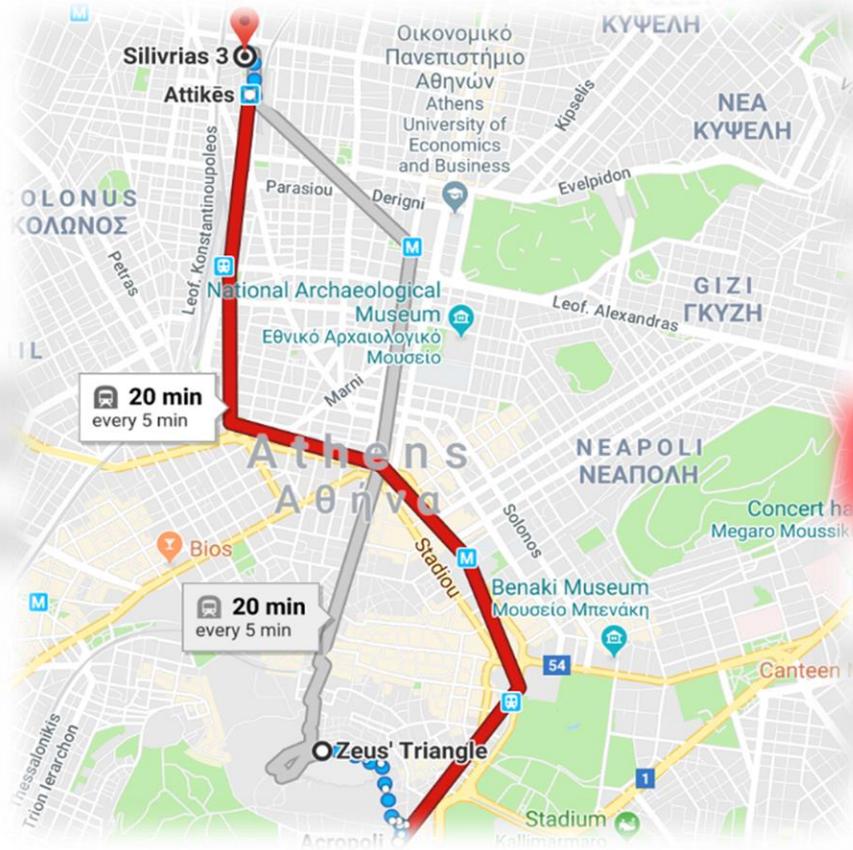
About distance (3/3)

Zeus' Triangle, Athina 105 58

Silivrias 3, Athina 104 45

Leave now

OPTIONS



Descriptive statistics:
Distance

4:45 PM–5:05 PM 20 min

M2

4:54 PM from Acropoli

11 min every 5 min

DETAILS

If Athens was one listing

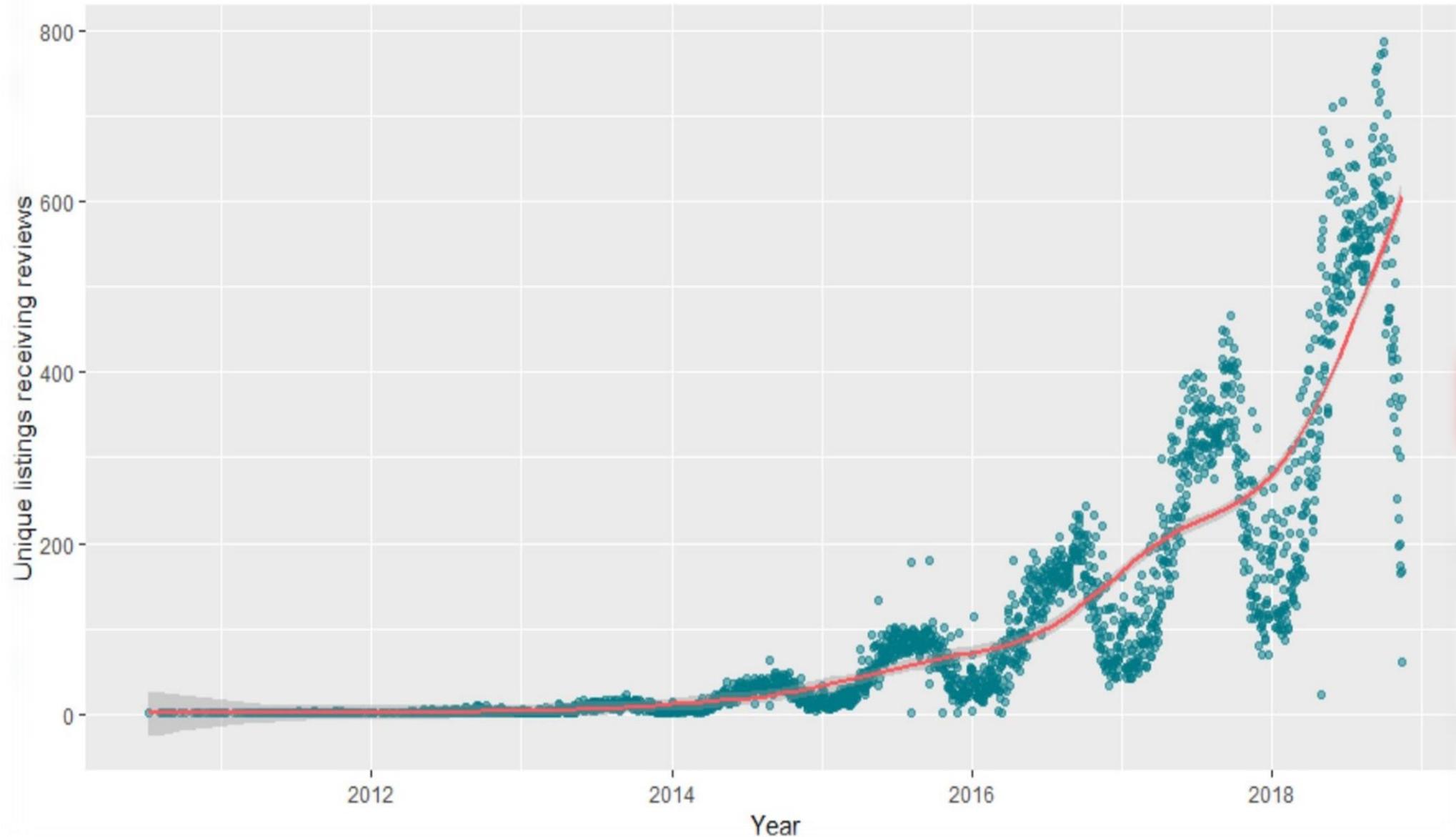


**Descriptive
statistics:
One Listing**

Avg. New Price 54.25 €	Median New Price 45.55 €
Avg. Walking Time 32.33 minutes	Avg. Transit Time 22.8 minutes
Avg. Reviews Score 95.29	Avg. Reviews/Month 2.52
Avg. <u>Nr.</u> of Listings/Host 11.48	Avg. People Capacity 3.93

How popular is Airbnb in Athens?

Number of Reviews across years

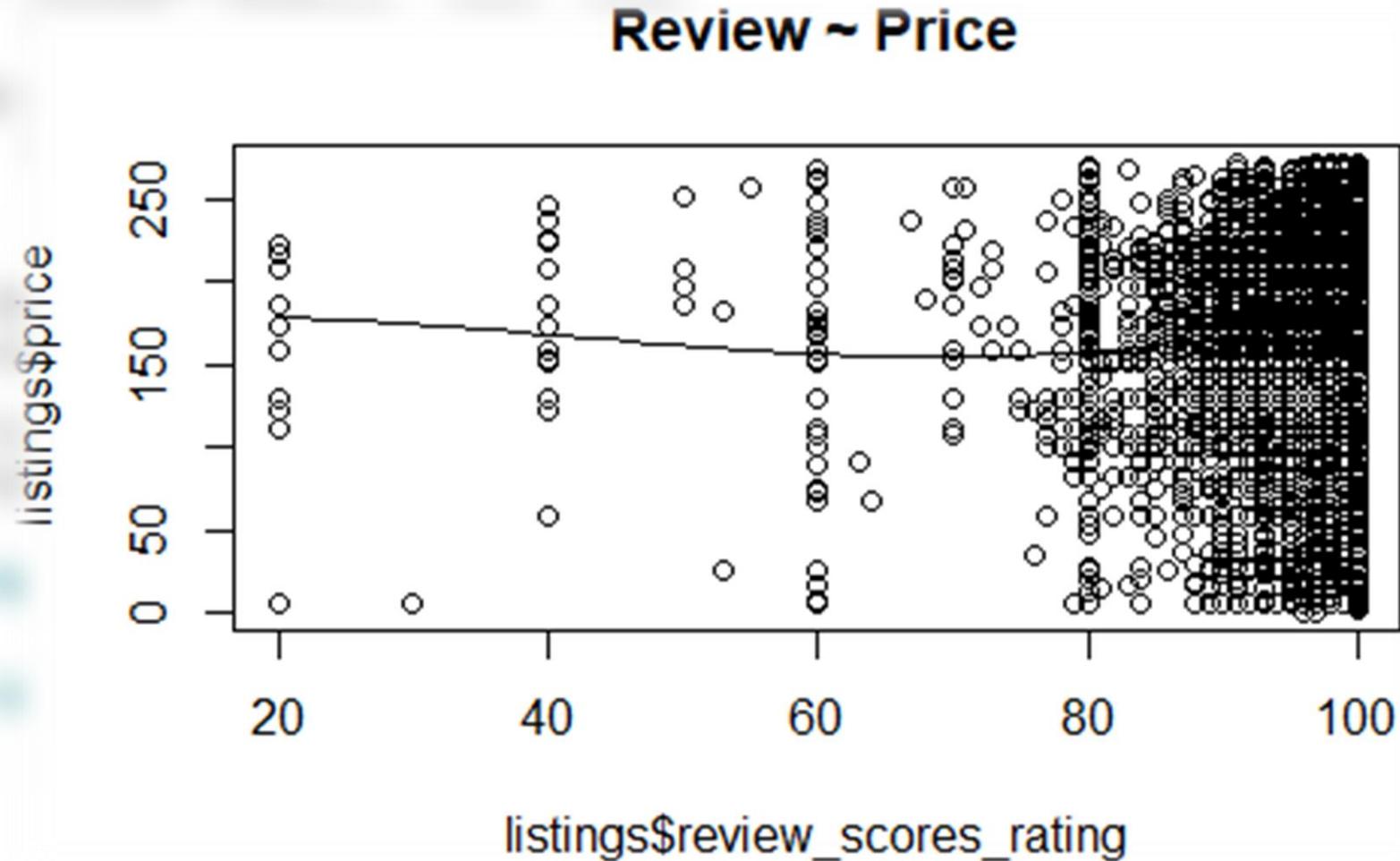


Reviews
(1/2)

The problem with reviews



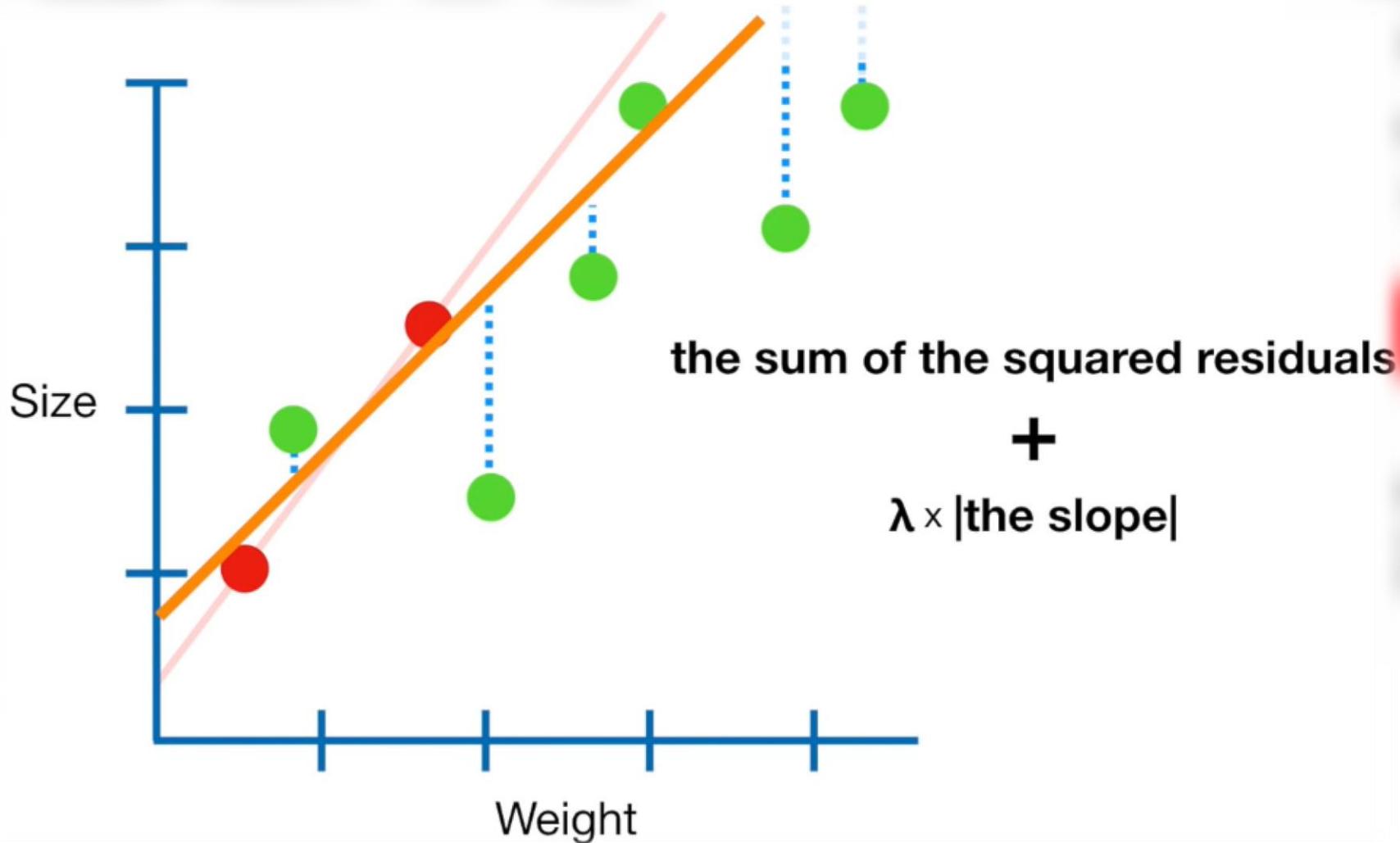
Reviews
(2/2)



Results



Model



What is LASSO
(Least Absolute
Shrinkage and
Selection
Operator)?

accommodates 3.6709406	bathrooms 18.85376551	bedrooms 3.681258421
email 0	phone 0	facebook 0
government_id 0	jumio 0	offline_government_id 0
identity_manual 0	reviews 0	work_email 0
kba 0	selfie 0	google 0
host_has_profile_pic 0	is_superhost 0	TV 3.729437624
Air.conditioning 4.752341756	Elevator 0	Kitchen 0
Family.kid.friendly 0	Parking 0	Wifi 0
Heating 0	Cooking 0	Patio.or.balcony 0
Washer 0	Smoking.allowed -2.17609324	Pets.Allowed 0
Smoke.detector 3.244869822	Hot.water 0	Luggage.dropoff.allowed 0
X24.hour.check.in 0	pool 33.53340307	Distance -6.503019394
review_scores_rating 0	minimum_nights 0	Days_since_host 0
host_total_listings_count 0		number_of_reviews 0



Lasso regression

Variables	Walking Time	Transit Time
accommodates	4.610897*** (28.53512)	4.826473*** (29.04444)
Air.conditioning	6.642961*** (8.650843)	6.714741*** (8.499663)
bathrooms	17.13866*** (0.672806)	17.59968*** (25.4353)
bedrooms	2.0841 (5.093552)	1.65061*** (3.924199)
Smoke.detector	3.61048 (7.162466)	4.171005*** (8.063573)
Smoking.allowed	-3.490986*** (-6.710173)	-3.617636*** (-6.756742)
TV	4.328316*** (7.506078)	3.960515*** (6.678178)
walking_minutes/ transit_minutes	-0.421693*** (-30.07475)	-0.780776*** (-24.26313)
pool	52.63587*** (19.94427)	51.88456*** (19.1123)
C	13.20141*** (11.57505)	16.73178*** (12.7265)



Robust Least Regression with Lasso
 “survivors for walking and transit time

Conclusions

- Host attributes and reviews attributes are penalized by lasso, so only traditional accommodation attributes and amenities remain
- Still, hosts seem to matter, but only during and after the stay, so this might be reflected on the reviews (the reviews problem)
- A lot more variables were not taken into account (**host description, host profile picture, rental rules**)
- Distance from subway stations, as well as distance from Acropolis is a noteworthy factor



Almost there,
hang on

But why?



Almost there,
hang on

The screenshot shows the landing page for 'Beyond Pricing' by 'Wheelhouse'. The page features a teal header with the 'eliot&me' logo on the left and navigation links for 'Property Managers', 'Vacation Rental Owners', 'How It Works', and 'Su'. Below the header, a secondary navigation bar includes 'Wheelhouse', 'Why Wheelhouse', 'How it works', 'Who it's for', 'What it costs', and 'Where we are'. The main content area has a light gray grid background. On the left, a vertical teal bar contains the text 'A p Pro Va S'. The central headline reads 'The smartest way to price your short-term rental'. Below this, a sub-headline states 'Earn 10-40% more revenue with data-driven pricing software built for all vacation rental hosts, owners, and property managers.' Two call-to-action buttons are present: a purple 'View the demo' button and a white 'Get started' button with a purple border. At the bottom, a footer line reads '30-day free trial · No card required · Cancel any time'. On the right side of the page, there is an illustration of a laptop with two callout bubbles: one showing '21% REVENUE LAST MONTH' and another showing '14% THIS WEEK'S OCCUPANCY'.

eliot&me

Beyond Pricing Property Managers Vacation Rental Owners How It Works Su

Wheelhouse Why Wheelhouse How it works Who it's for What it costs Where we are

The smartest way to price your short-term rental

Earn 10-40% more revenue with data-driven pricing software built for all vacation rental hosts, owners, and property managers.

[View the demo](#) [Get started](#)

30-day free trial · No card required · Cancel any time

21% REVENUE LAST MONTH

14% THIS WEEK'S OCCUPANCY



Thank you for your attention!



**Final Slide,
yeah**

**Have questions? Send them this way
gkove@okthess.gr**

The Lasso regression



The Model

$$\sum_{i=1}^n \left(y_i - \sum_j x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

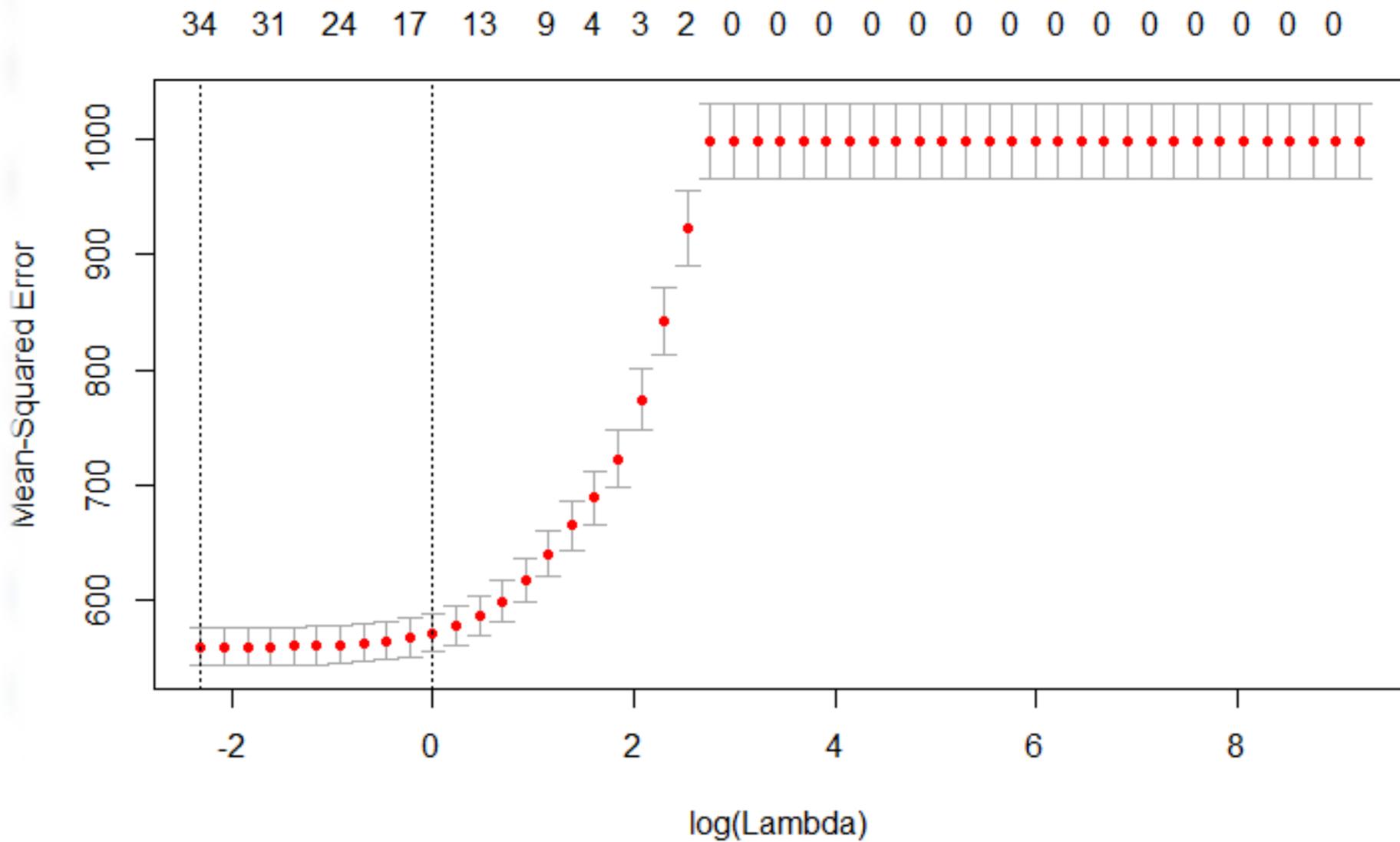
λ = tuning parameter

επιλέγουμε

$$\lambda = 10^i \text{ (για } i = -1, -0.9, \dots, 4)$$



Lasso regression





accommodates	3.677603
Air.conditioning	3.551086
bathrooms	18.410304
bedrooms	3.310354
Distance	-5.820369
pool	27.172566
Smoke.detector	2.066279
Smoking.allowed	-1.578161
TV	2.767273
C	18.936869
R-squared	0.3993866
Mean Estimate of standard error	26.758

2nd Lasso regression with lasso “survivors”

accommodates		3.657206***	
		(16.60287)	
Air.conditioning		7.132***	
		(6.816957)	
bathrooms		21.02328***	
		(22.93896)	
bedrooms		4.869805***	
		(8.731773)	
Distance		-7.232125***	
		(-27.0113)	
pool		49.88444***	
		(13.87649)	
Smoke.detector		4.701728***	
		(6.844518)	
Smoking.allowed		-4.416073***	
		(-6.231795)	
TV		6.090611***	
		(7.749145)	
C		10.33213	
		(6.903197)	
R-squared	0.412354	Mean	54.26205
		dependent var	
Adjusted R-squared	0.411585	S.D.	31.60165
		dependent var	
S.E. of regression	24.24106	Akaike info criterion	9.215425



OLS Regression with Lasso survivors

accommodates	4.642744*** (28.40484)
Air.conditioning	6.677408*** (8.600772)
bathrooms	17.53849*** (25.78987)
bedrooms	2.015254*** (4.869726)
Distance	-5.518362*** (-27.77625)
pool	50.93833*** (19.09602)
Smoke.detector	3.626719*** (7.115134)
Smoking.allowed	-3.633168*** (-6.909495)
TV	4.444983*** (7.621612)
C	9.06826*** (8.165228)



**Robust Least
Squares
Regression with
Lasso survivors**



Robust Statistics			
R-squared	0.263092	Adjusted R-squared	0.262127
Rw-squared	0.528961	Adjust Rw-squared	0.528961
Akaike info criterion	8599.316	Schwarz criterion	8671.245
Deviance	2303256	Scale	16.38154
Rn-squared statistic	7024.167	Prob(Rn-squared stat.)	0

Non-robust Statistics			
Mean dependent var	54.26205	S.D. dependent var	31.60165
S.E. of regression	24.77376	Sum squared resid	4218228

Robust Least Regression statistics with Lasso survivors



Robust Least Regression with Lasso
 “survivors for walking and transit time

Robust Statistics			
R-squared	0.269254/ 0.249860	Adjusted R-squared	0.268298/ 0.248878
Rw-squared	0.53715/ 0.516586	Adjust Rw-squared	0.53715/ 0.516586
Akaike info criterion	8693.569/ 8627.972	Schwarz criterion	8765.354/ 8699.772
Non-robust Statistics			
Mean dependent var	54.26205/ 54.26205	S.D. dependent var	31.60165/ 31.60165
S.E. of regression	24.63342/ 25.12405	Sum squared resid	4170573/ 4338360.

accommodates	7.2642
bathrooms	9.3366
bedrooms	8.4464
identity_manual	5.2605
TV	8.35506
Air.conditioning	12.6827
Elevator	-2.5354
Kitchen	2.9621
Smoking.allowed	-4.1911
Smoke.detector	2.3522
pool	8.9917
distance	-3.0361
R-Squared	0.4618
Mean Estimate of standard error	41.4764



Robust Least Regression for Rome with Lasso “survivors”

Discussion

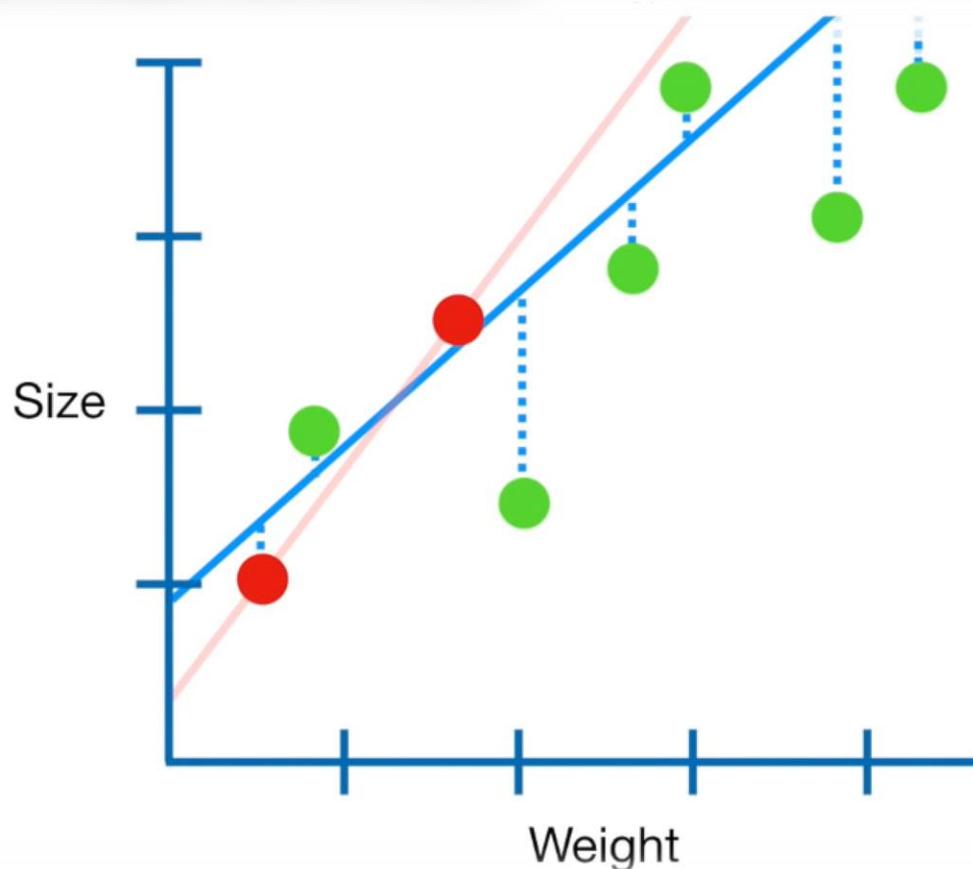
- Research on neighborhood attributes could help “catch” neighborhood effects and premia
- Also, research on rentals and hotel industry impact could be useful for policy-makers
- Airbnb seems to be in Athens to stay



**Final Slide,
yeah**

Appendix

- What is LASSO (Least Absolute Shrinkage and Selection Operator)?



Final Slide,
yeah