

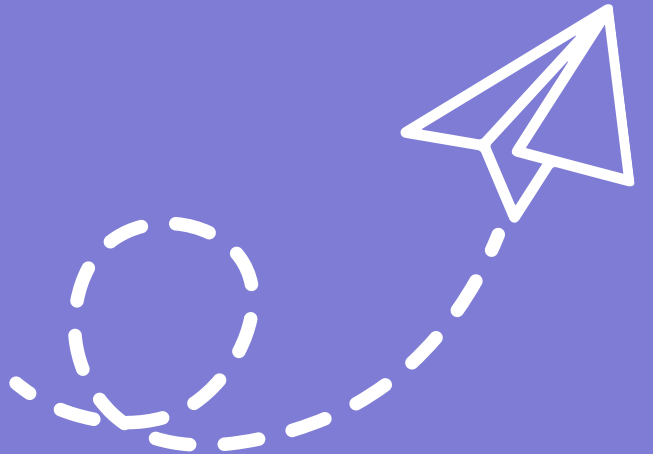

MIDDLE-BOOTCAMP PROJECT  
DATA ANALYTICS IRONHACK  
ROCIO ALVAREZ-REMENTERIA

# EXPECTATIONS VS REALITY

MADRID HOUSE MARKET

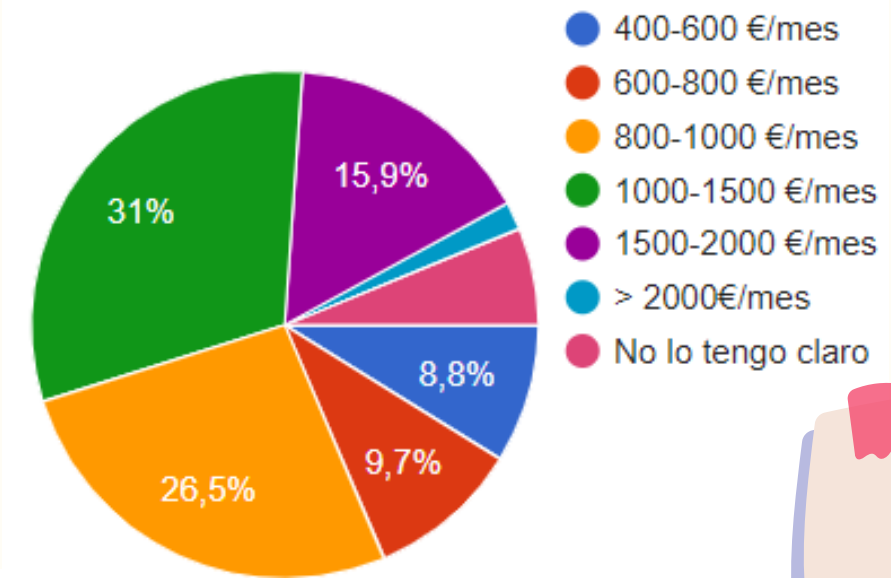
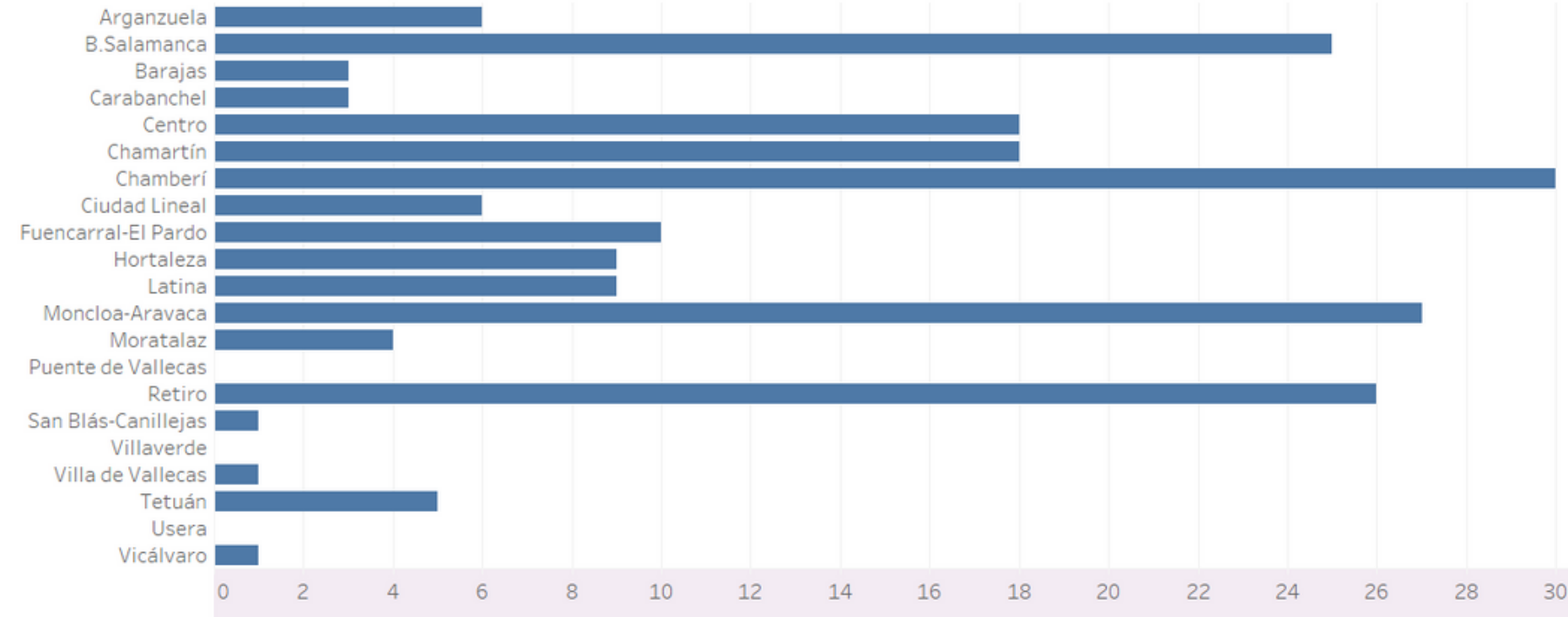


# *Challenges*

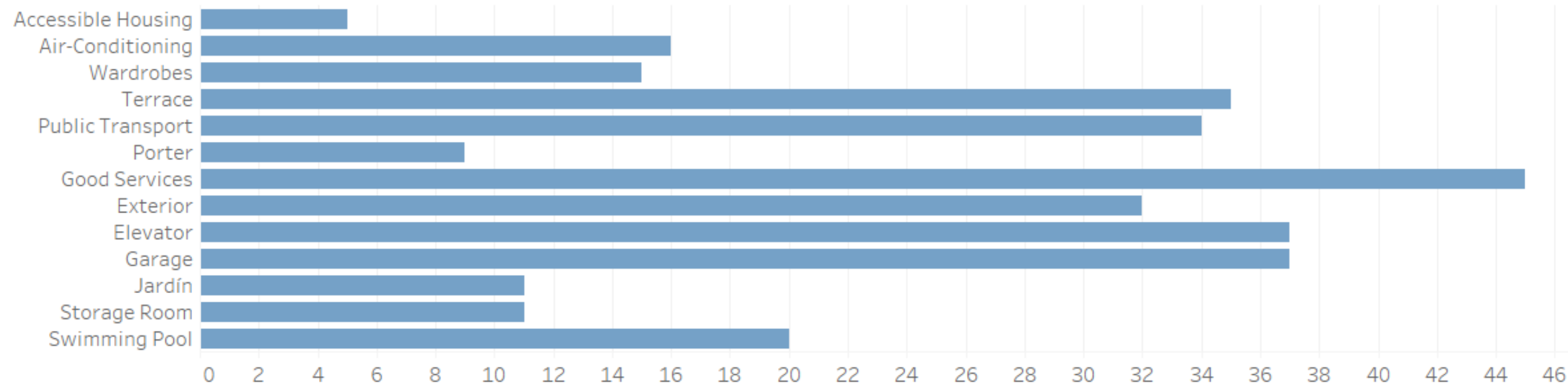
1. Survey analysis & transformation
  2. Estimation of mortgage and savings
  3. Outliers detection & bucketization process
  4. Web-scrapping at 'El Idealista'
  5. Market data analysis
  6. Matches between clients and houses
  7. Expectations analysis
- 
- 

# SURVEY

## Districts



## Client priorities



Survey:  
117 records

# MORTGAGE SIMULATION

```
In [57]: def max_mor(row):
        quota = row['max_quota']
        interest = 4
        years = row['max_years']

        # Annual rate:
        monthly_interest = interest / 12 / 100

        # Quotas
        total_quotas = years * 12

        # Amortization factor
        amortizacion_factor = (monthly_interest * (1 + monthly_interest) ** total_quotas) / ((1 + monthly_interest) ** total_quotas)

        # Total Loan
        total = quota / amortizacion_factor

        return total

        # Aplicar la función a cada fila del DataFrame
        data['max_loan'] = 1.2*(data.apply(max_mor, axis=1))
        data['max_saving'] = 0.4*(data.apply(max_mor, axis=1))

        # Mostrar el DataFrame con la nueva columna 'total_loan'
        data.head()
```

```
Out[57]:
```

ty_of_living	rooms	...	mortgage_years	house_area	min_quota	max_quota	min_years	max_years	min_loan	min_saving	max_loan	max_saving
partner	2	...	20-25	70-90	800.0	1000.0	20.0	25.0	158420.983914	52806.994638	227342.979570	75780.993190
partner	2	...	30-35	70-90	600.0	800.0	30.0	35.0	150812.093127	50270.697709	216814.535933	72271.511878
partner	3	...	20	90-120	800.0	1000.0	20.0	NaN	158420.983914	52806.994638	NaN	NaN
partner	3	...	25-30	90-120	400.0	600.0	25.0	30.0	90937.191828	30312.397276	150812.093127	50270.697709
partner	3	...	30-35	90-120	1000.0	1500.0	30.0	35.0	251353.488545	83784.496182	406527.254874	135509.084958

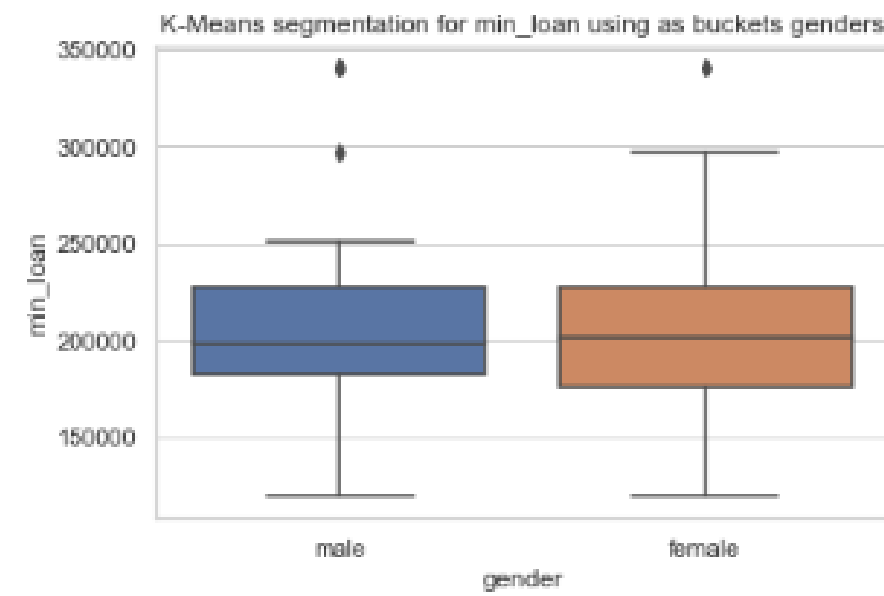
4% of annual rate  
Max quota, max years  
40% of savings

# DATA TRANSFORMATION

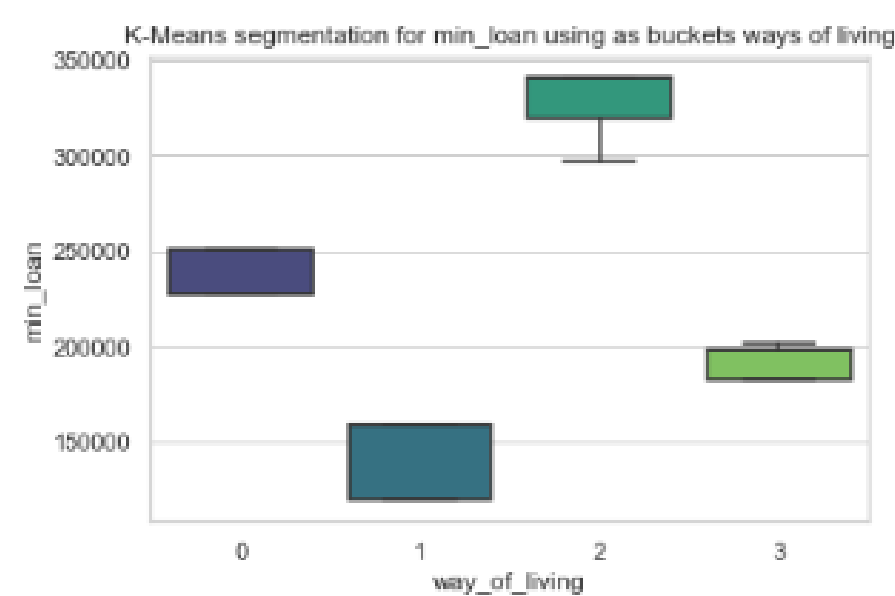
¿Preferirías vivir de alquiler o compra?	¿Cuántos años tienes?	¿Cuál es tu género?	Selecciona los distritos dónde buscarías actualmente casa (se pueden marcar hasta 5 opciones):	¿En qué estado prefieres la vivienda? (puedes seleccionar varias opciones)	¿Qué aspectos son para ti más importantes a la hora de buscar vivienda? (puedes marcar hasta 5 opciones)	¿En qué planta te gustaría vivir?	¿Prefieres vivir solo o acompañado?	¿Cuántas habitaciones buscas?	¿Cuántos baños necesitas?	¿Cuánto dinero al mes estarías dispuesto a destinar (entre tú y las personas con las que vivirías para pagar la vivienda si decidieras convivir)?	Si compraras una casa, ¿durante cuántos años estarías dispuesto a hipotecarte?
Compra	Entre 25 y 30	Masculino	Fuencarral-El Pardo	En buen estado	Buenos servicios (supermercados, cafeterías, f...	Plantas intermedias, Casa/Chalet	Con mi pareja	2	1	800-1000 €/mes	entre 20 y 25
Compra	Entre 30 y 40	Masculino	Hortaleza	A reformar	Buenos servicios (supermercados, cafeterías, f...	Plantas intermedias	Con mi pareja	2	1	600-800 €/mes	entre 30 y 35
Compra	Entre 30 y 40	Femenino	Centro, Retiro, B.Salamanca	Nueva	Buen transporte público, Terraza, Piscina, Bue...	Ático	Con mi pareja	3	2	800-1000 €/mes	menos de 20
Compra	Entre 25 y 30	Femenino	Retiro, B.Salamanca, Fuencarral-El Pardo, Hort...	En buen estado	Buen transporte público, Ascensor, Garaje, Bue...	Ático, Casa/Chalet	Con mi pareja	3	2	400-600 €/mes	entre 25 y 30
Compra	Entre 30 y 40	Femenino	Chamberí, Moncloa-Aravaca, Ciudad Lineal	En buen estado	Buen transporte público, Terraza, Garaje, Buen...	Plantas intermedias	Con mi pareja	3	2	1000-1500 €/mes	entre 30 y 35

	districts	max_loan	min_area	garage	rooms	swimming_pool	elevator	porter	terrace	public_transport	air_conditioning	storage_
	fuencarral	227342.979570	70.0	0	2	0	0	0	0	0	0	
	hortaleza	216814.535933	70.0	0	2	0	0	0	0	0	0	
	mberí,moncloa_aravaca,ciudad_lineal	406527.254874	90.0	1	3	0	0	0	1	1	0	
	centro,hortaleza,canillejas,barajas	227342.979570	NaN	1	3	0	1	0	0	1	1	

# POBLATION ANALYSIS

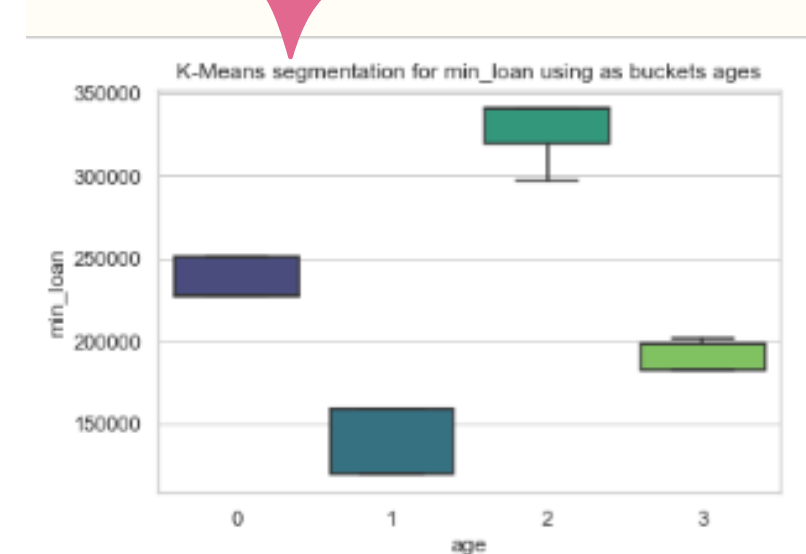


**Male vs. female Risk Appetite**  
p-value = 0.87



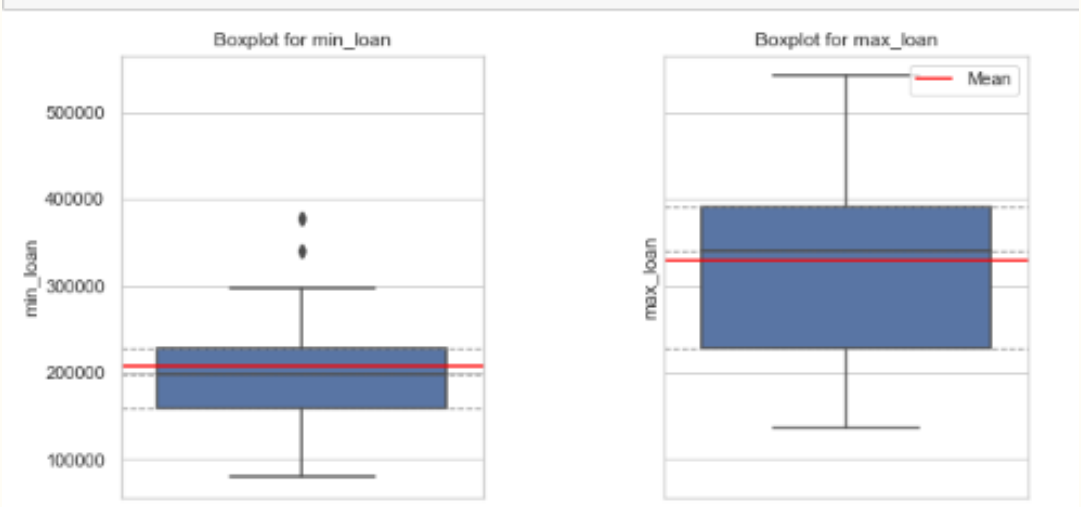
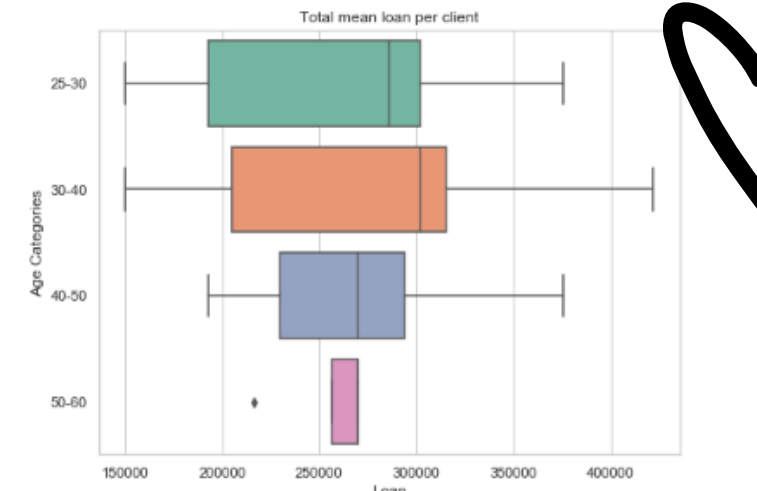
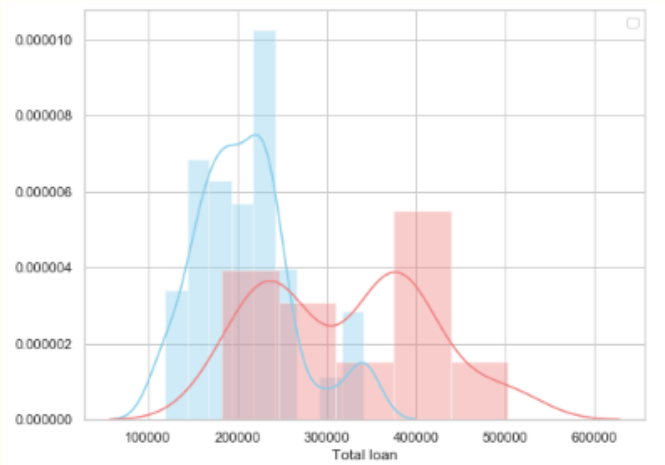
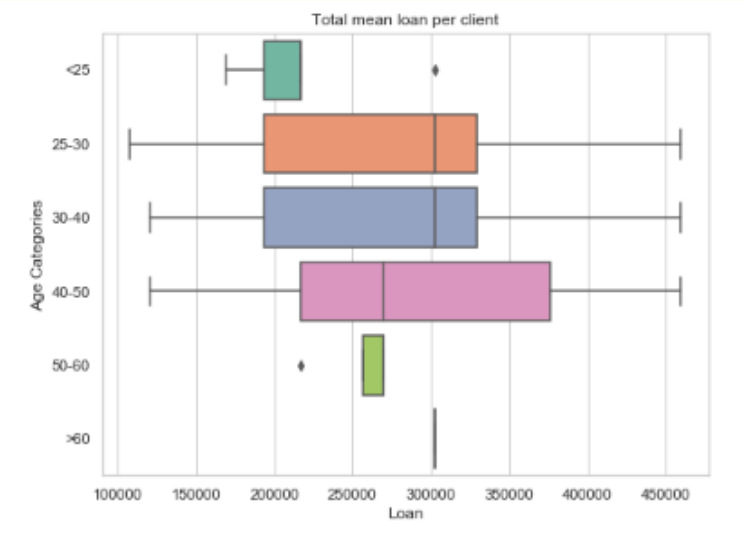
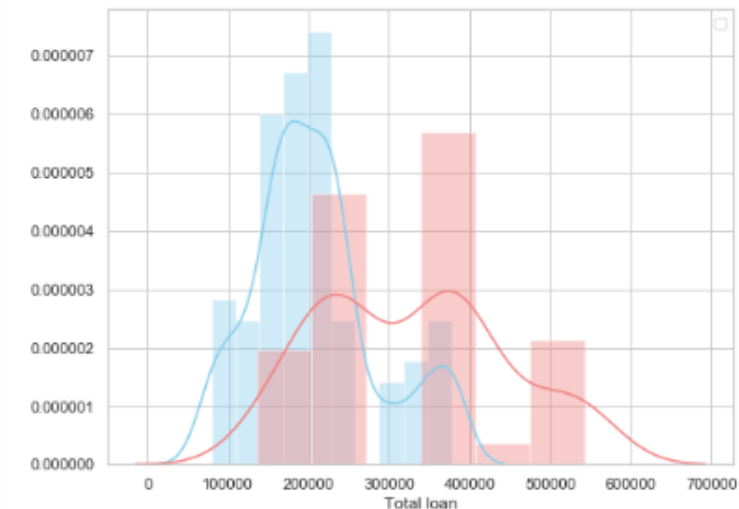
**Ways of living Risk Appetite:**

1. Alone
  2. With family
  3. With your partner
  4. With friends
- p-value =  $3.19 \times 10^{-31}$

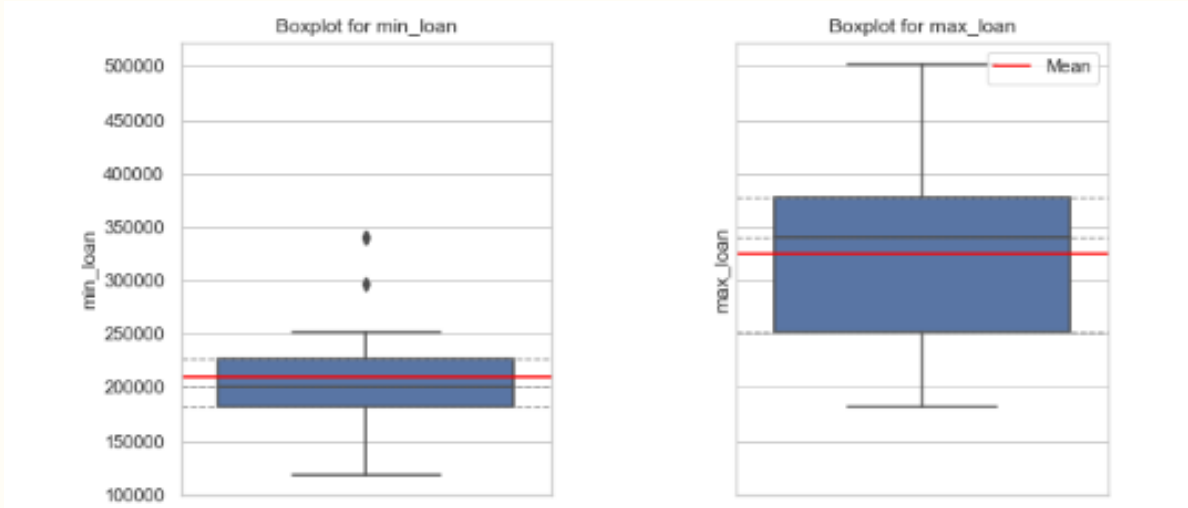


**Age factor - Risk Appetite**  
p-value =  $5.48 \times 10^{-41}$

# OUTLIERS DETECTION

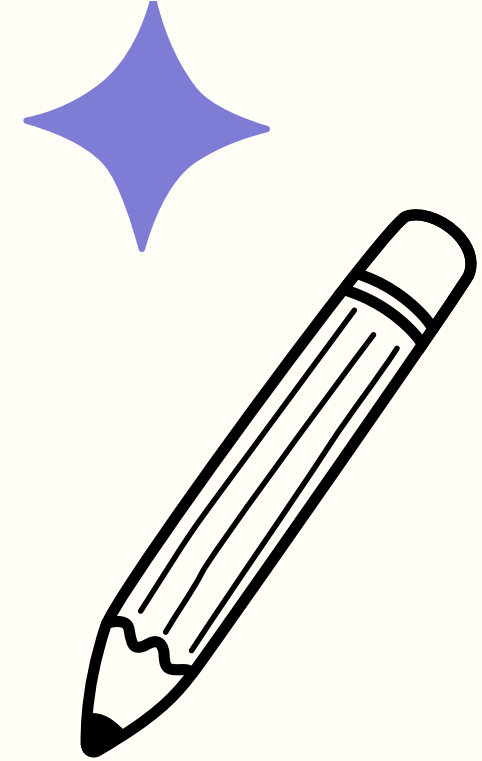


Before outliers detection



After outliers detection

Quantile 90%  
Discard criteria for Age >60





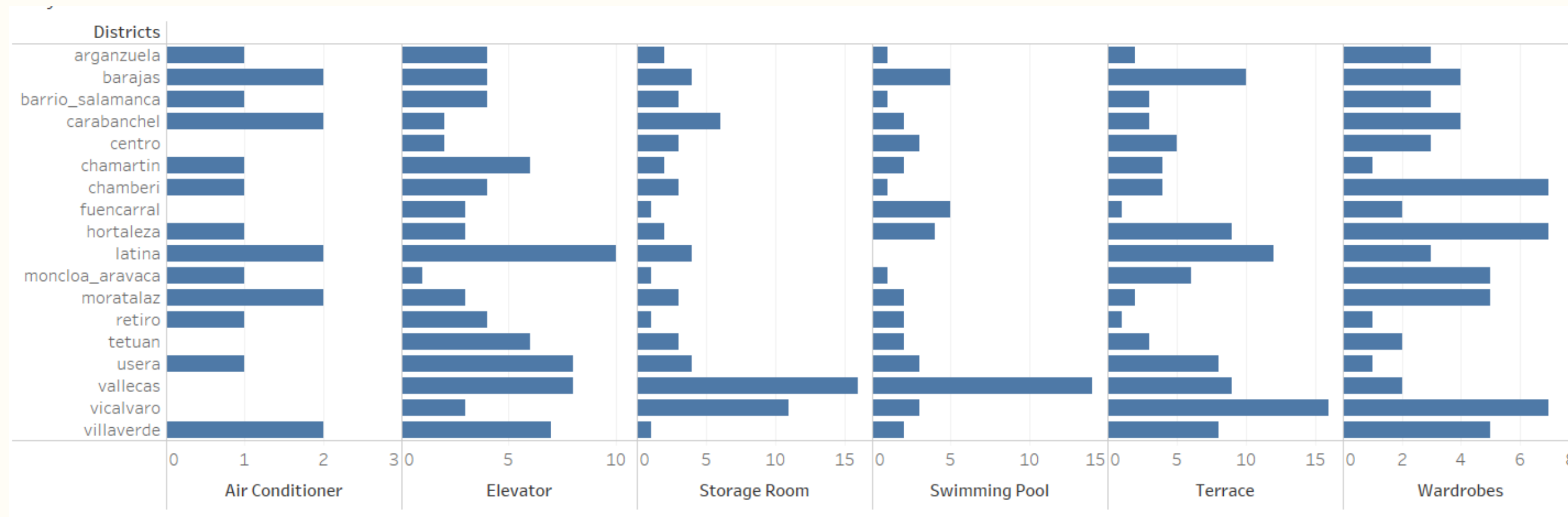
# MARKET PRICE ANALYSIS

Web scrapping  
at  
El Idealista:  
628 records

chamartin	arganzuela	centro	moncloa_aravaca	barajas	vicalvaro
barrio_salamanca	hortaleza				
tetuan	chamberi	vallecas		latina	carabanchel
fuencarral	retiro	villaverde			
		moratalaz		usera	



# MARKET ATTRIBUTES ANALYSIS



# MATCHES AT BARRIO DE SALAMANCA

*3 matches vs. 25 records*

```
len_max_loan = len(exp_barrio_salamanca['max_loan'])
len_price = len(mkt_barrio_salamanca['price'])

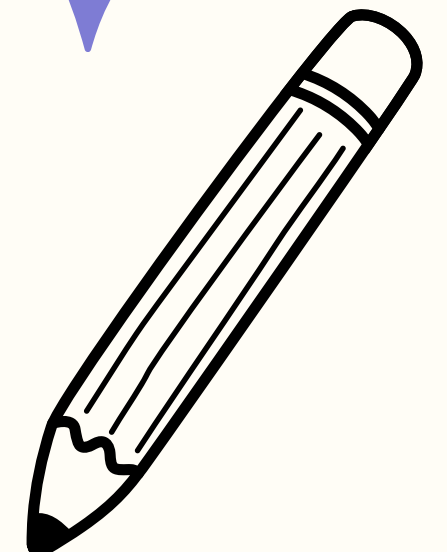
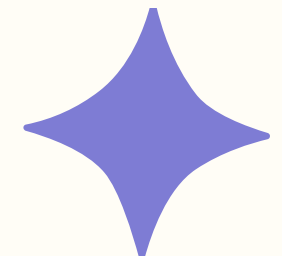
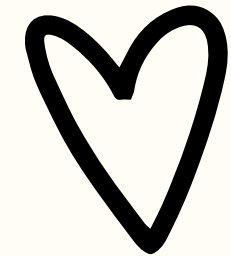
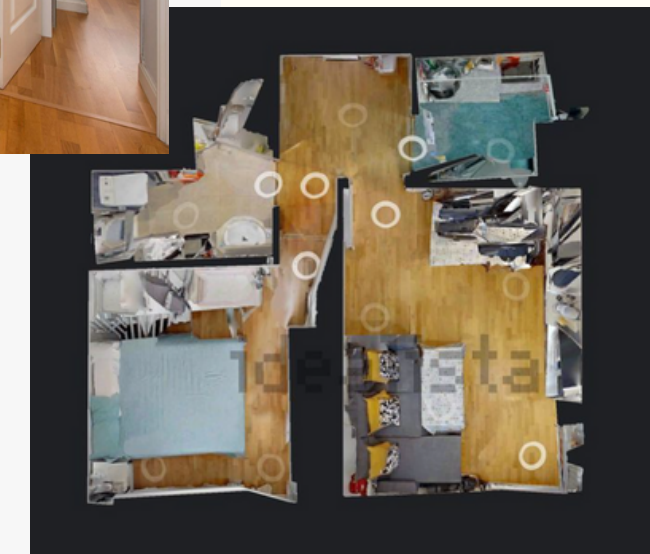
if len_max_loan != len_price:

    min_len = min(len_max_loan, len_price)

    exp_barrio_salamanca = exp_barrio_salamanca.head(min_len)
    mkt_barrio_salamanca = mkt_barrio_salamanca.head(min_len)

filtered_mkt_barrio_salamanca = mkt_barrio_salamanca[
    (exp_barrio_salamanca['max_loan'].values >= mkt_barrio_salamanca['price'].values) &
    (exp_barrio_salamanca['elevator'].values == mkt_barrio_salamanca['elevator'].values) &
    (exp_barrio_salamanca['terrace'].values == mkt_barrio_salamanca['terrace'].values)
]

filtered_mkt_barrio_salamanca[['title', 'price', 'swimming_pool', 'terrace', 'air_conditioner', 'exterior', 'wardrobes', 'garden', 'stor
```



	title	price	swimming_pool	terrace	air_conditioner	exterior	wardrobes	garden	storage_room	elevator
331	Piso en Secoya 3 D, Pau de Carabanchel, Madrid	380000	1	0	0	0	0	0	1	0
332	Piso en calle Ardemans, Guindalera, Madrid	297000	0	0	0	0	0	0	0	0
337	Piso en calle de Ardemans, Guindalera, Madrid	299000	0	0	0	0	0	0	0	0
346	Piso en Roma, Guindalera, Madrid	350000	0	1	0	0	0	0	0	0

# MATCHES AT CHAMBERÍ

*1 match vs. 30 records*

In [142]:

```
len_max_loan = len(exp_chamberi['max_loan'])
len_price = len(mkt_chamberi['price'])
```

```
if len_max_loan != len_price:
```

```
    min_len = min(len_max_loan, len_price)
```

```
    exp_chamberi = exp_chamberi.head(min_len)
```

```
    mkt_chamberi = mkt_chamberi.head(min_len)
```

```
filtered_mkt_chamberi = mkt_chamberi[
    (exp_chamberi['max_loan'].values >= mkt_chamberi['price'].values) &
    (exp_chamberi['elevator'].values == mkt_chamberi['elevator'].values) &
    (exp_chamberi['terrace'].values == mkt_chamberi['terrace'].values)
]
```

```
filtered_mkt_chamberi[['title', 'price', 'swimming_pool', 'terrace', 'air_conditioner', 'exterior', 'wardrobes', 'garden', 'storage_room', 'elevator']]
```

Out[142]:

	title	price	swimming_pool	terrace	air_conditioner	exterior	wardrobes	garden	storage_room	elevator
394	Ático en calle de Fernando el Católico, Gaztam...	250000	0	0	0	0	0	0	0	0



# MATCHES AT MONCLOA:

3 match vs. 27 records

```
len_max_loan = len(exp_moncloa_aravaca['max_loan'])
len_price = len(mkt_moncloa_aravaca['price'])

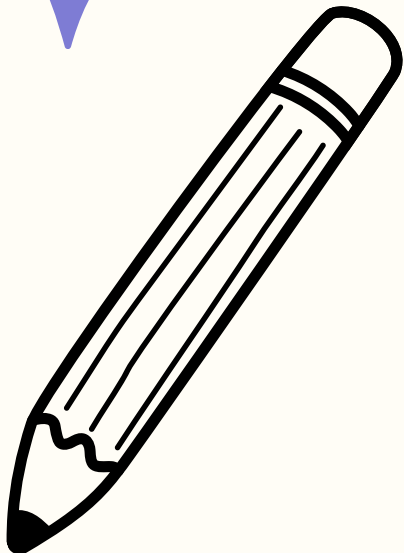
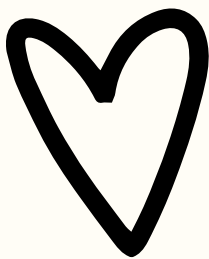
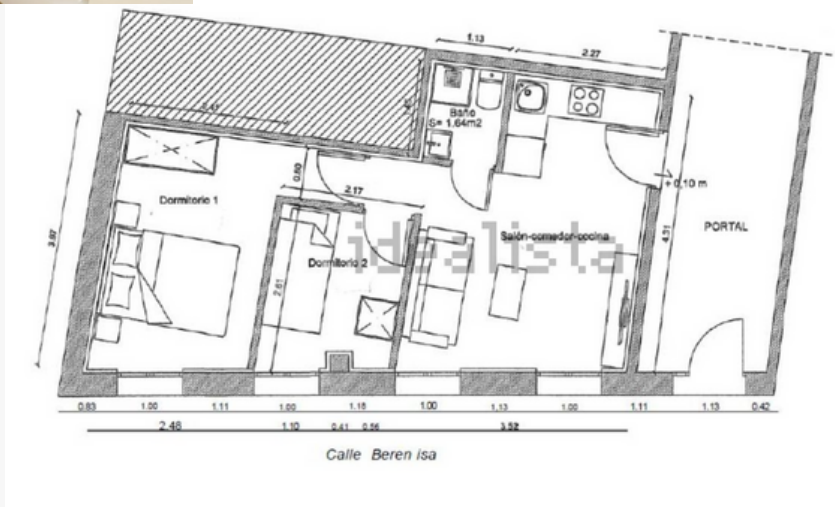
if len_max_loan != len_price:
    min_len = min(len_max_loan, len_price)

    exp_moncloa_aravaca = exp_moncloa_aravaca.head(min_len)
    mkt_moncloa_aravaca = mkt_moncloa_aravaca.head(min_len)

filtered_mkt_moncloa_aravaca = mkt_moncloa_aravaca[
    (exp_moncloa_aravaca['max_loan'].values >= mkt_moncloa_aravaca['price'].values) &
    (exp_moncloa_aravaca['elevator'].values == mkt_moncloa_aravaca['elevator'].values) &
    (exp_moncloa_aravaca['terrace'].values == mkt_moncloa_aravaca['terrace'].values)
]

filtered_mkt_moncloa_aravaca[['title', 'price', 'swimming_pool', 'terrace', 'air_conditioner', 'exterior', 'wardrobes', 'garden', 'storage_room', 'elevator']]
```

	title	price	swimming_pool	terrace	air_conditioner	exterior	wardrobes	garden	storage_room	elevator
422	Piso en Berenisa, Aravaca, Madrid	250000	0	0	0	0	0	0	0	0
431	Casa o chalet independiente en Guadarrama	399000	0	0	0	0	0	0	0	0
435	Piso en Valdemarín, Madrid	277000	0	0	0	0	0	0	0	0



# MATCHES

	market price	expectations price	market vs. expectations
B. Salamanca	382.044.00 €	374.173.00 €	2%
Chamberí	373.863.64 €	350.693.00 €	6%
Moncloa/Aravaca	430.000.00 €	349.058.00 €	19%

***B.Salamanca: 3 matches vs. 25 records***

***Chamberi :1 match vs. 30 records***

***Moncloa/Aravaca: 3 matches vs. 2 records***

## Can be possible:

- Terrace
- Good services

## Over expectations:

- Wardrobes
- Exterior
- Air-conditioning
- Garden
- Swimming pool
- Elevator



**THANKS!**

