

# Residential Building Data Set

*Projet de Modèle régression régularisées*

---

MRR

ENSIIE

Groupe 9  
Guangyue CHEN  
Jiahui XU

---

---

# Introduction

---

- ❖ Le dataset compose le coût de construction, les prix de vente, les variables de projet et les variables économiques correspondant aux appartements résidentiels et immobiliers situés à Téhéran, en Iran.
- ❖ La dimension de data: 372 observations 109 variable
- ❖ Deux sorties: Le prix de vente réels et le coûts de construction réels(Actual sales prices & Actual construction costs)

---

# L'IDÉE

---

- ❖ Construire deux modèles pour les deux sorties. En considérant que les deux sorties (les objectifs) et les attributs sont continus, nous nous décidons utiliser Linear Regression.



---

# La Méthodologie

---

- ❖ Pour sélectionner le modèle, nous utilisons la mesure “L’erreur Quadratique Moyenne” (MSE) et la méthode ‘K-folds’ Validation croisée. Nous comparons les performances des modèles sur le nouvelle échantillon, a la fin nous choisissons la modèle ‘Elastic Net’ pour le objectif ‘Le prix de vente réels’, et la modèle ‘Ridge’ pour le objectif ‘le coûts de construction réels’.

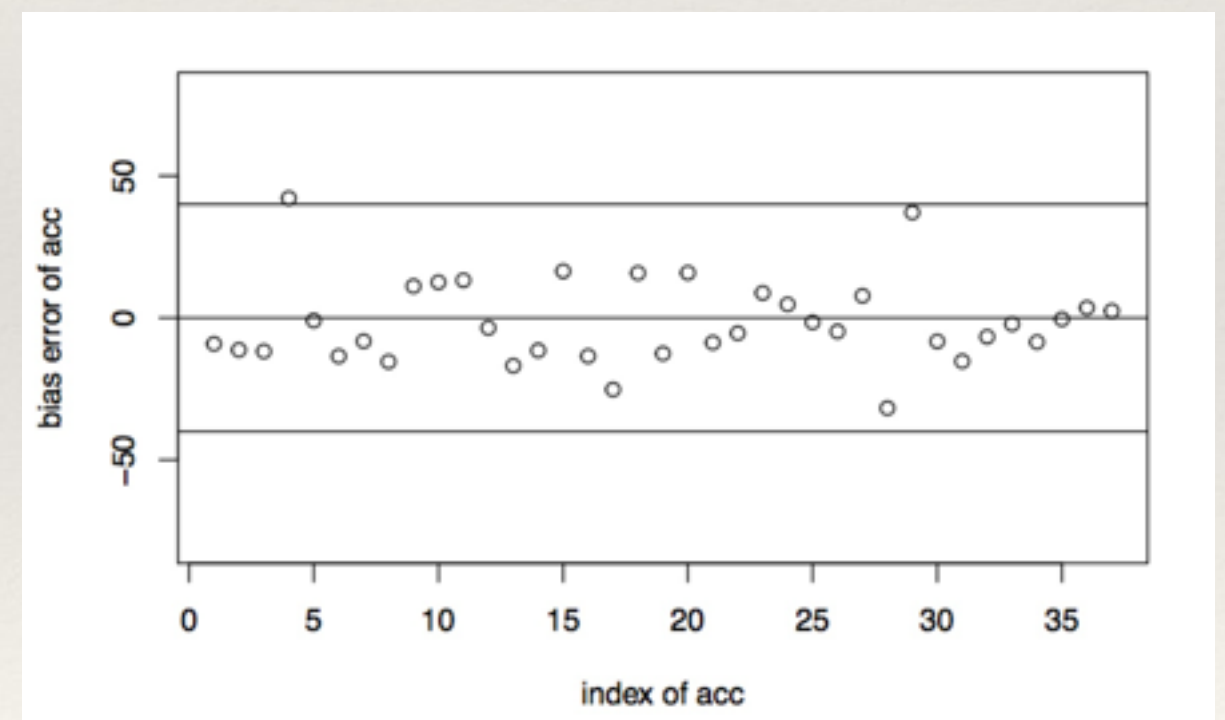
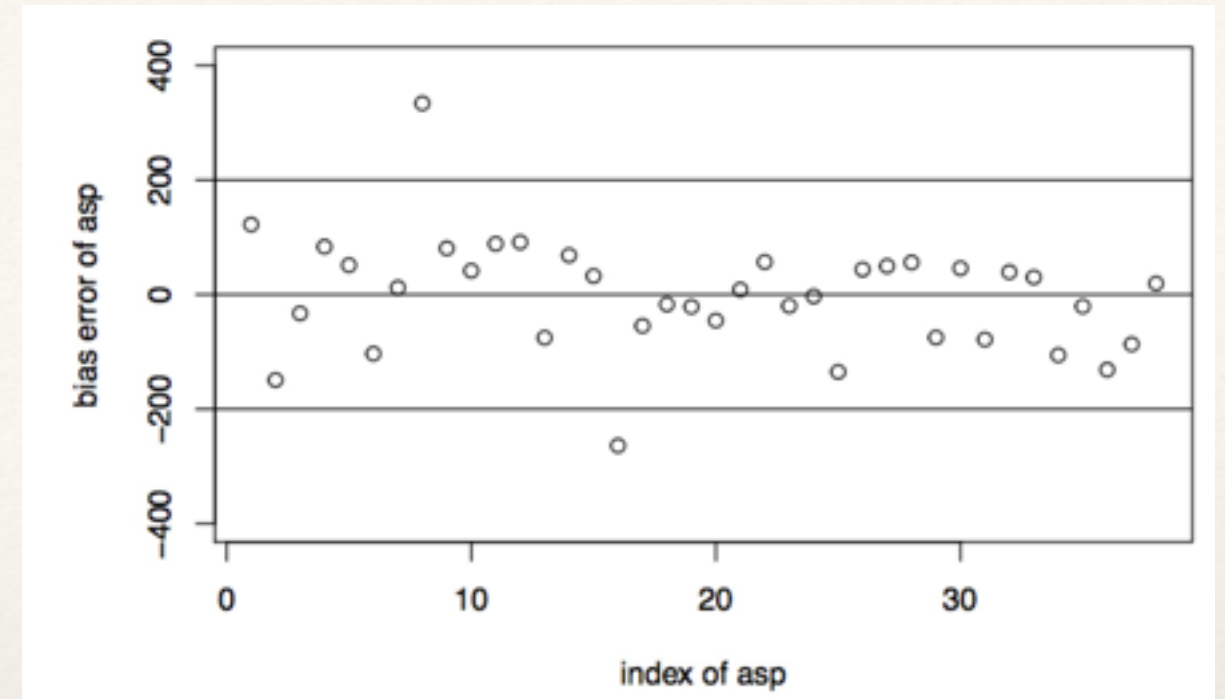
---

# Tester les modèle

---

Les résultats du modèle comporteront les erreurs de biais.

Pour "Prix de vente réels", la prédiction est égale à la valeur réelle  $\pm 200$  et pour "Coûts de construction réels", la prévision est égale à la valeur réelle  $\pm 40$ .

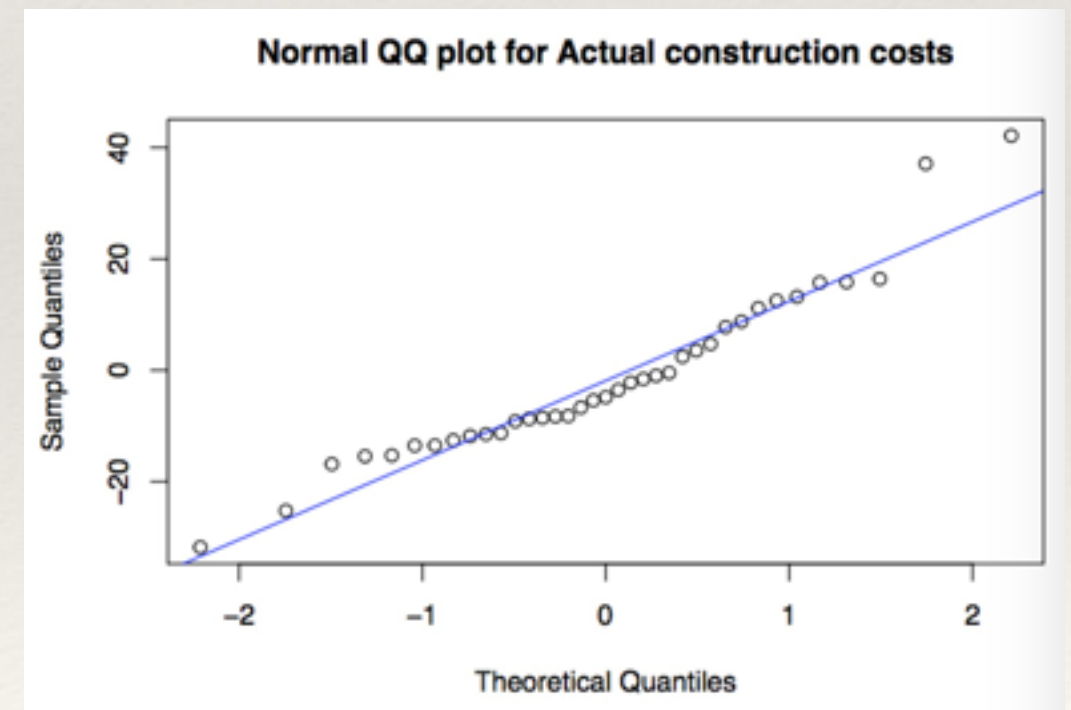
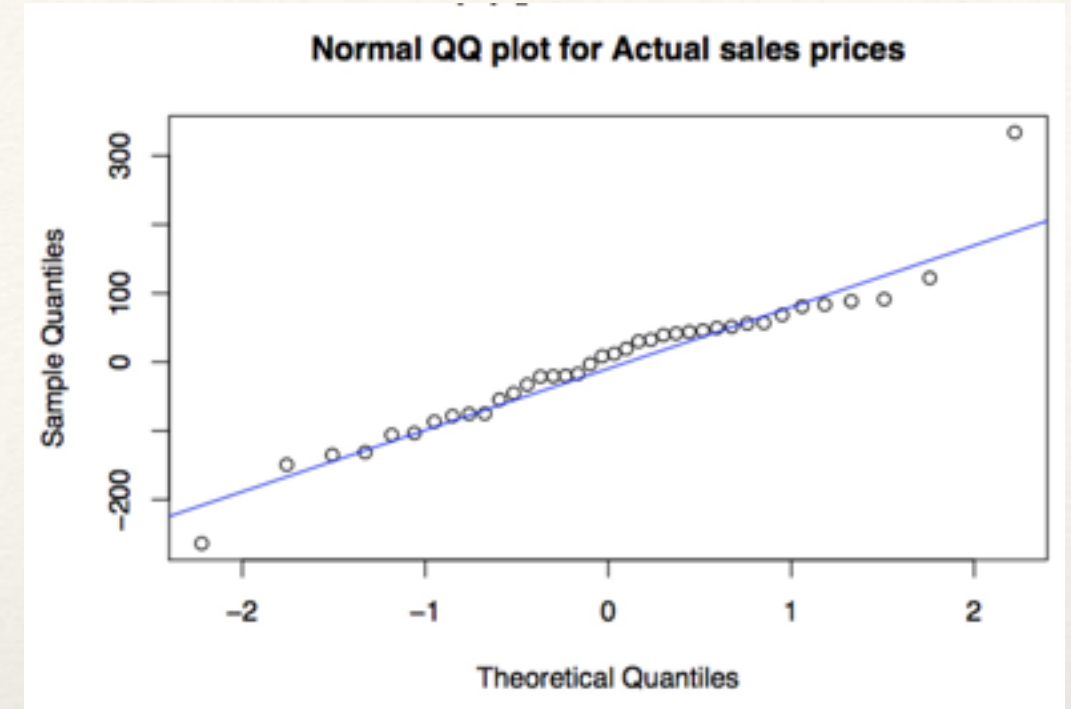


---

# Tester les modèle

---

Donc nous voyons que le tracé Q-Q normal est presque une ligne droite. Nous pouvons donc résumer que ce sont de bons modèles.





---

# La résultât

---

- ❖ A la fin , nous choisissons le lambda qui minimise les biais de nouveaux data. En utilisant ce lambda , nous entraînons tous les data pour obtenir la modèle finale.
- ❖ la modèle 'Elastic Net' pour le objectif 'Le prix de vente réels'
- ❖ la modèle 'Ridge' pour le objectif 'le coûts de construction réels'

---

# Le prix de vente réels

---

- ❖ les coefficients:
- ❖ completion quarter:17.5
- ❖ Project locality defined in terms of zip codes:-5.29
- ❖ Lot area:-0.15
- ❖ Preliminary estimated construction cost based on the prices at the beginning of the project:-0.45
- ❖ Duration of construction: 35.70
- ❖ Price of the unit at the beginning of the project per m2: 1.18
- ❖ Total floor areas of building permits issued by the city municipality:-5.32 22.51 -5.15 3.34
- ❖ The interest rate for loan in a time resolution:-0.09 25.27 -17.83 11.06
- ❖ The number of loans extended by banks in a time resolution: 0.11 0.19
- ❖ Wholesale price index (WPI) c of building materials for the base year:-0.10
- ❖ Land price index for the base year: -7.09



---

# Le coûts de construction réels

---

- ❖ les coefficients:
- ❖ Start year:3.93
- ❖ Start quarter:-43.73
- ❖ Complete year:3.93
- ❖ completion quarter:1.23
- ❖ Project locality defined in terms of zip codes:-0.30
- ❖ Preliminary estimated construction cost based on the prices at the beginning of the project:1.41
- ❖ Duration of construction: 13.28
- ❖ Building services index (BSI) b for a preselected base year :-0.11 -0.27 -0.19 -0.29 -0.59
- ❖ Wholesale price index (WPI) c of building materials for the base year:-0.23 -0.62 -0.65 0.76
- ❖ Total floor areas of building permits issued by the city / municipality:10.51 -3.71 -0.34 0.78 -0.80
- ❖ Land price index for the base year:0.20 0.26 -0.29
- ❖ The number of loans extended by banks in a time resolution :0.22 -0.15 0.18
- ❖ The interest rate for loan in a time resolution:-7.01 6.86 -1.01 7.54 -12.99
- ❖ The average construction cost of buildings by private sector at the time of completion of construction:0.10
- ❖ Consumer price index (CPI) i in the base year:0.38 0.18 -0.19 -0.25 -0.27
- ❖ CPI of housing, water, fuel & power in the base year: 0.12 0.14 -0.10 -0.12

---

# La Conclusion

---

- ❖ La régression avec régularisation convient très bien aux données, qui ont une dimension élevée et une multicollinéarité.
- ❖ Donc nous voyons que pour la sortie différente, les variables importantes sont différentes.



---

# Le prix de vente réels après normalisation

---

- ❖ les coefficients:
- ❖ completion quarter:19.68
- ❖ Project locality defined in terms of zip codes:-34.72
- ❖ Total floor area of the building: 82.01
- ❖ Lot area:-76.80
- ❖ Total preliminary estimated construction cost based on the prices at the beginning of the project: 10.28
- ❖ Preliminary estimated construction cost based on the prices at the beginning of the project:-50.81
- ❖ Equivalent preliminary estimated construction cost based on the prices at the beginning of the project in a selected base year:12.23
- ❖ Duration of construction: 74.83
- ❖ Price of the unit at the beginning of the project per m2: 1184.02