

# Modelos de regresión aplicados a una base de datos de la provincia de Teherán, Irán

Alejandro Garcia Carballo  
Juan Plúa Gutiérrez



# Introducción

1. Analizar los atributos de una base de datos y seleccionar los más representativos
2. Aplicar modelos de regresión
3. Analizar errores y mejorarlos con distintas técnicas
4. Resolución sobre un problema
5. Conclusiones



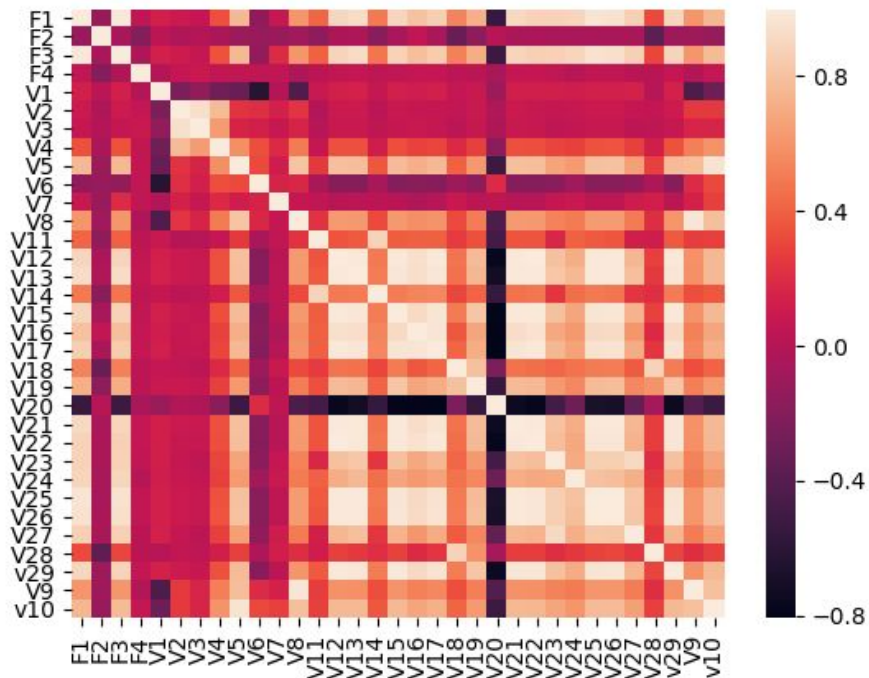
# Datos de entrada

Podemos diferenciar 3 bloques de datos:

- Fechas de proyectos
- Variables físicas y financieras del proyecto
- Variables económicas



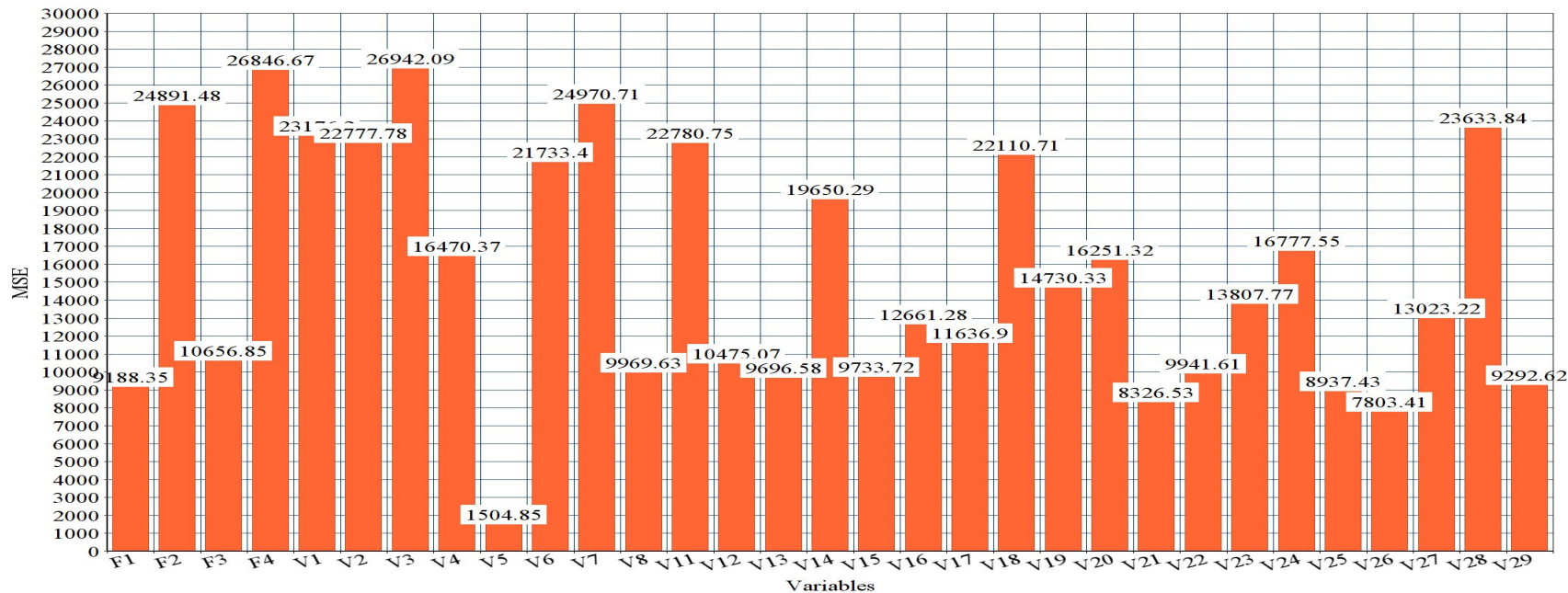
# Metodologia





# Metodologia

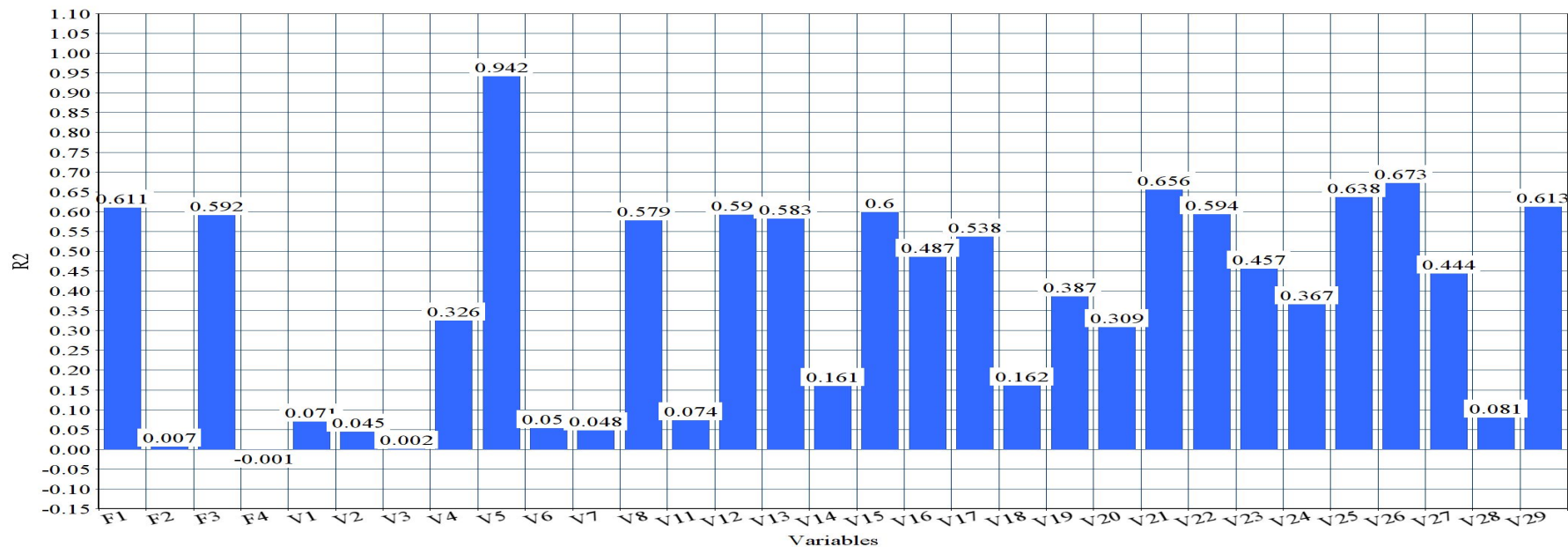
Comparación MSE variables individuales





# Metodología

Comparación R<sup>2</sup> variables individuales



# Desarrollo de un prototipo



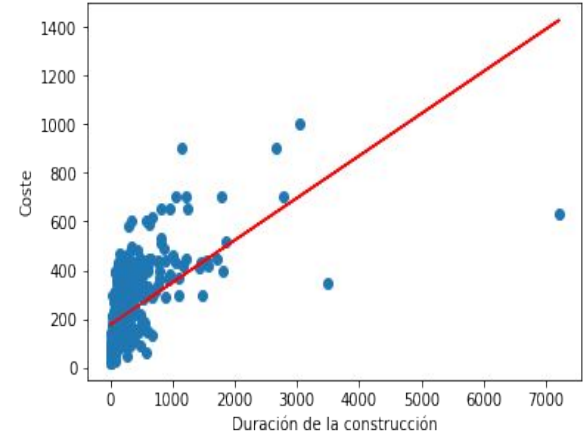
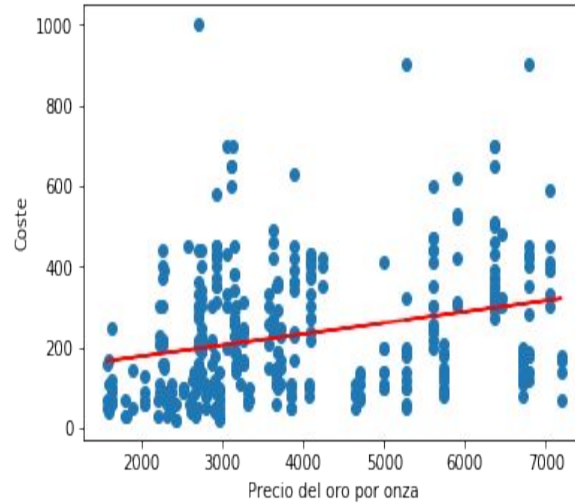
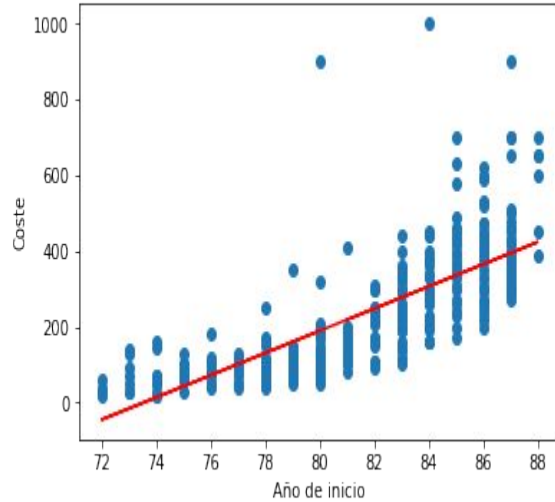


# Desarrollo de un prototipo

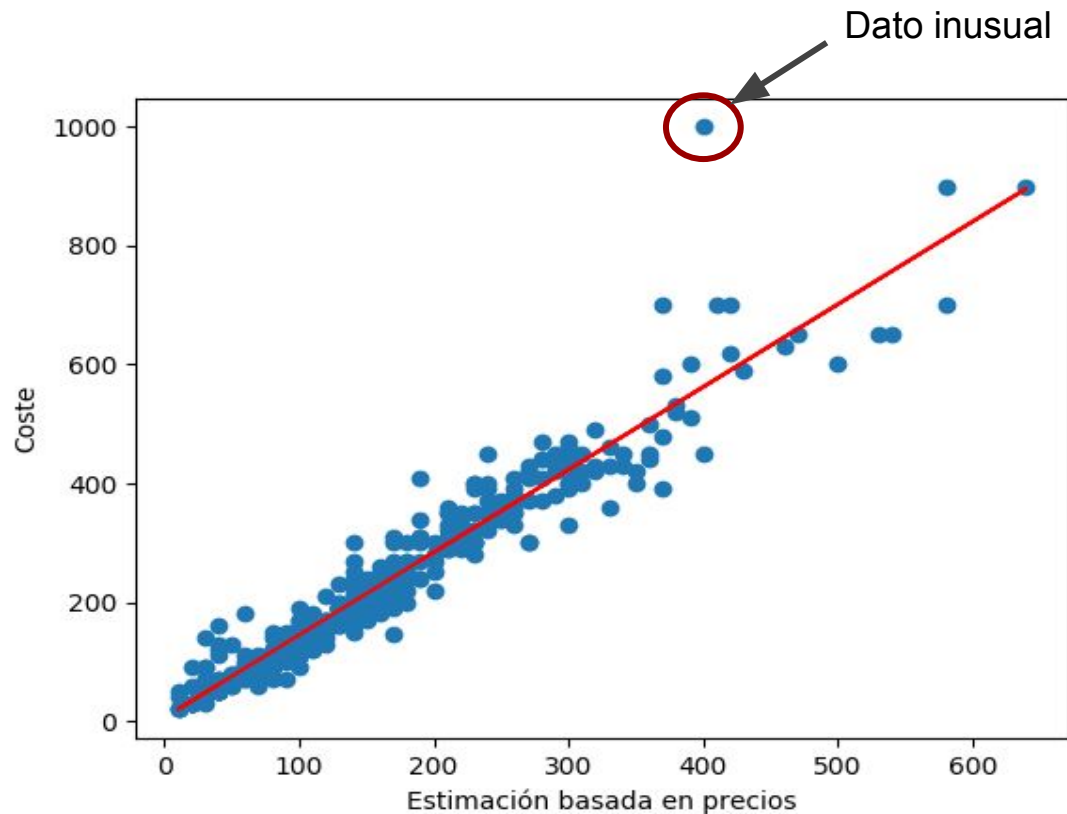
- Numpy
- Math
- Sklearn
- Pandas
- Matplotlib



# Desarrollo de un prototipo

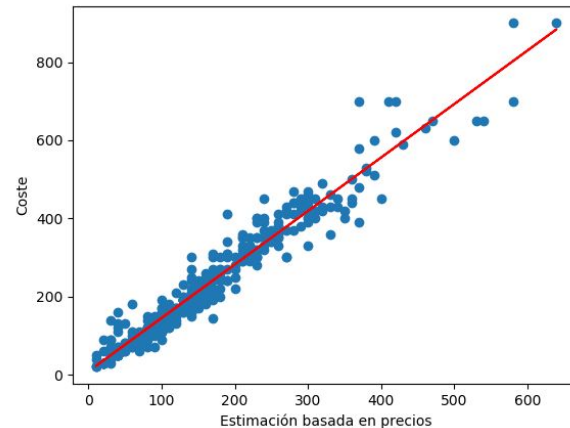
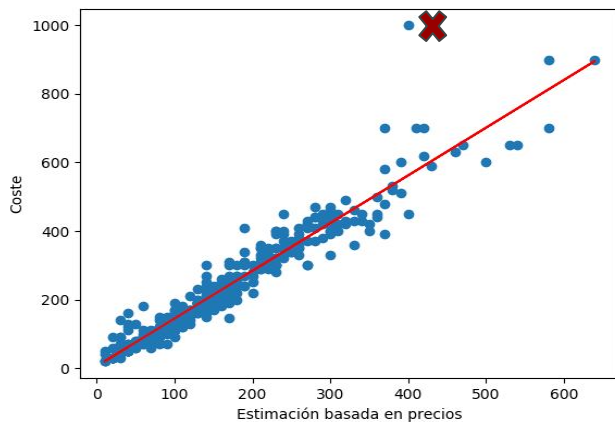


# Desarrollo de un prototipo





# Desarrollo de un prototipo



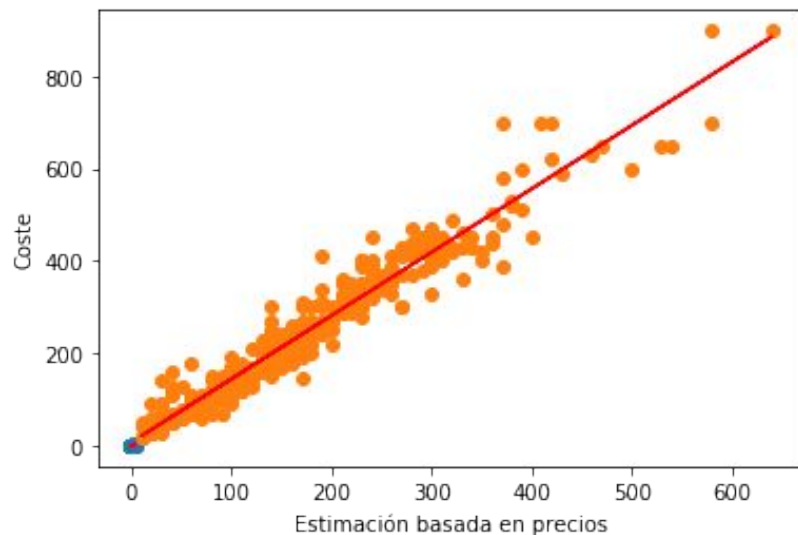


# Desarrollo de un prototipo

|                                 | <b>MSE</b> | <b>Root MSE</b> | <b>R2</b> |
|---------------------------------|------------|-----------------|-----------|
| <b>Todos los datos</b>          | 2093,30    | 45,75           | 0,92      |
| <b>Quitando datos inusuales</b> | 1279,26    | 35,76           | 0,94      |



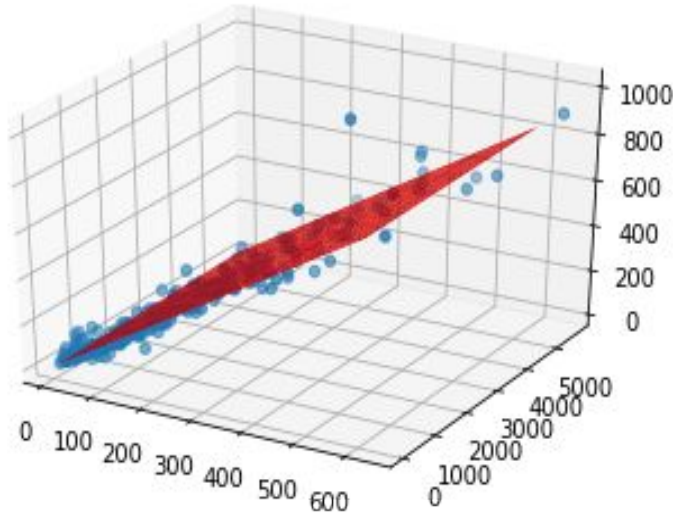
# Desarrollo de un prototipo



| MSE   | Root MSE | R2    |
|-------|----------|-------|
| 0.056 | 0.237    | 0.942 |

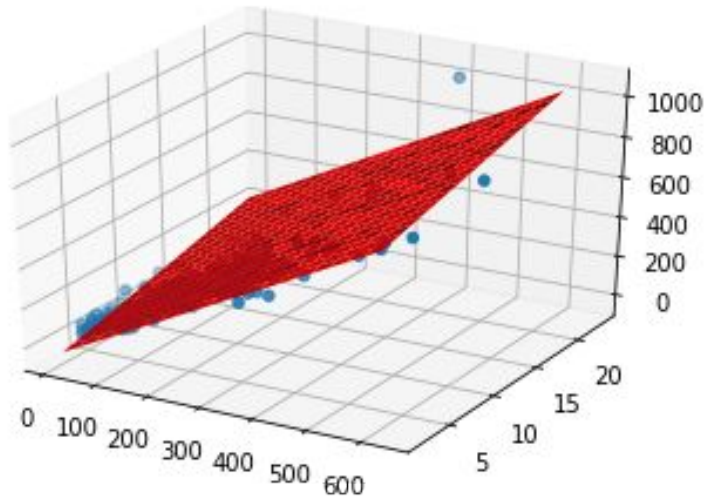


# Desarrollo de un prototipo



|                | <b>MSE</b> | <b>Root MSE</b> | <b>R2</b> |
|----------------|------------|-----------------|-----------|
| <b>V5 y V8</b> | 2075,49    | 45,56           | 0,92      |
| <b>V5</b>      | 1279,26    | 35,77           | 0,94      |

# Desarrollo de un prototipo



Variables V5 y V12 aumentan  $R^2$  al ser  
normalizadas (0,92  $\rightarrow$  0,94)



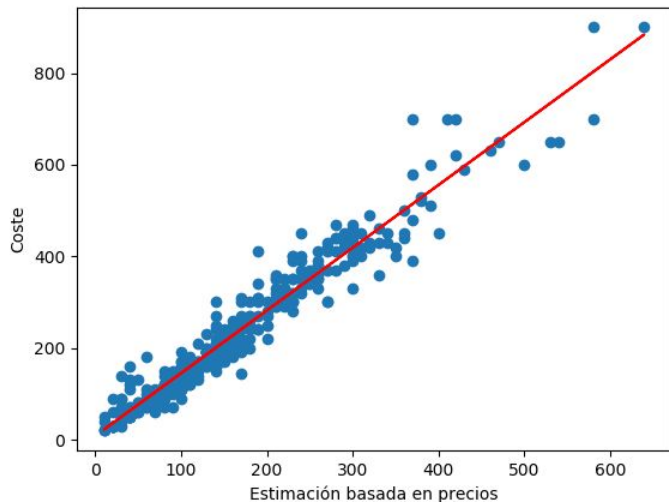
# Resultados del aprendizaje

|                    | <b>MSE</b> | <b>Root MSE</b> | <b>R2</b> |
|--------------------|------------|-----------------|-----------|
| <b>50% T 50% V</b> | 1612,7118  | 40,1558         | 0,9438    |
| <b>80% T 20% V</b> | 1279,2619  | 35,7667         | 0,9449    |





# Resultados del aprendizaje



| MSE     | Root MSE | R2    |
|---------|----------|-------|
| 1279,26 | 35,76    | 0,944 |



# Conclusiones sobre la viabilidad del aprendizaje

- Dificultades
  - Determinar si los resultados estaban siendo correctos
  - Mejorar nuestro modelo individual
- Conclusiones
  - Hemos conseguido un modelo con MSE bajo y  $R^2$  alto
  - Estimación del coste correcta



# Actividad

- El trabajo en grupo ha facilitado llevar a cabo el proyecto.
- Comprobación si al tener otros datos se podrían obtener mejores resultados con el mismo modelo.
- El aspecto más positivo ha sido aprender a aplicar diferentes técnicas para realizar un modelo de regresión.
- Durante la realización de la práctica hemos notado carencias a la hora de la validación de resultados.

# ¡Gracias por vuestra atención!

Alejandro Garcia Carballo  
Juan Plúa Gutiérrez