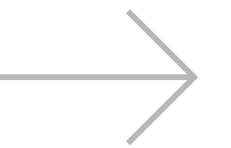
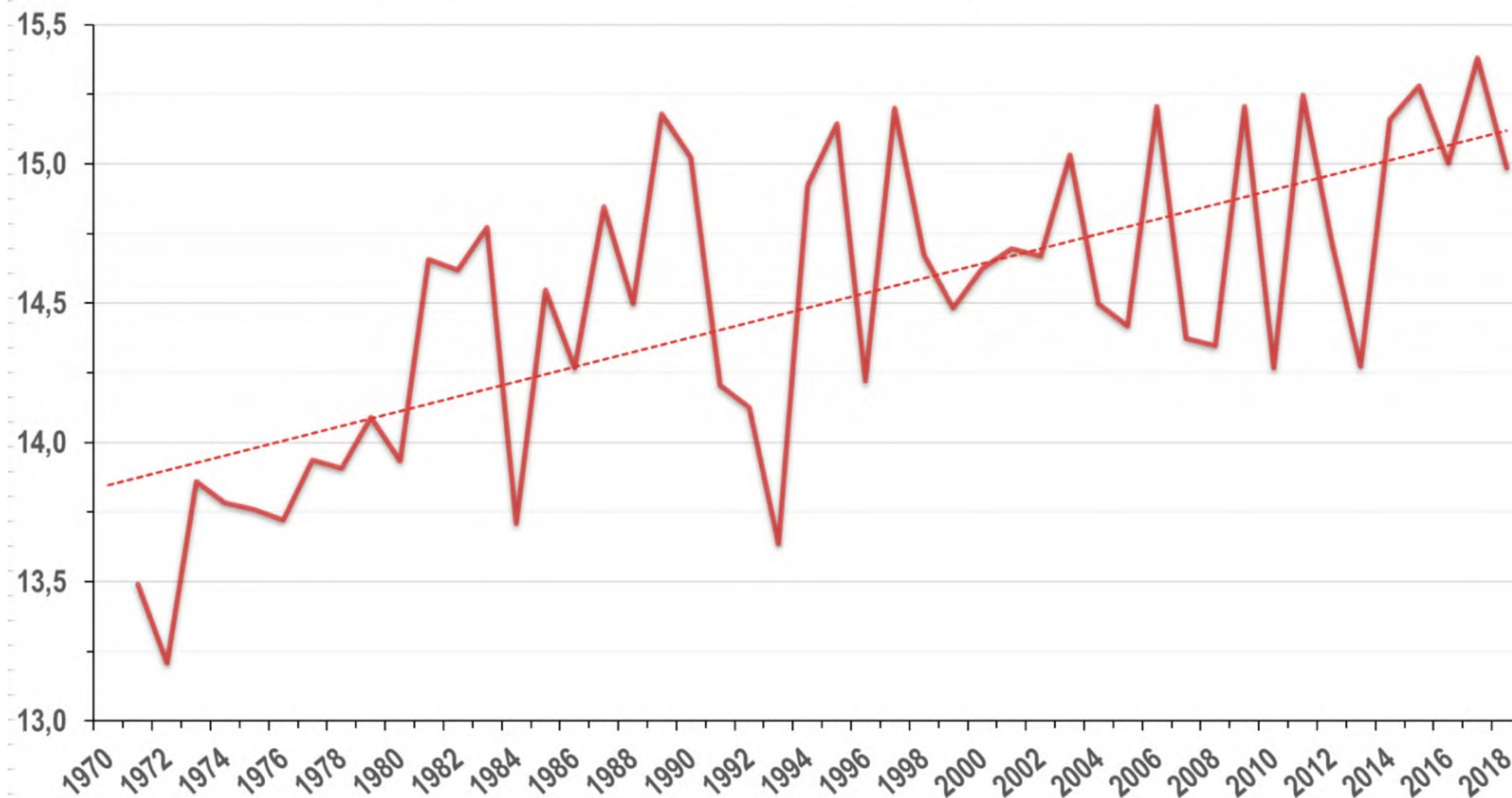


¿**Rioja** denominación de origen Antártida?

Data Analytics — Ironhack 2024



Evolución de la temperatura media anual (°C) de España 1971-2018



Objetivo

Predecir el impacto del
cambio climático en el vino



Creación del set de datos

01

Vinos y sus atributos

- Kaggle data set
- 100646 vinos únicos
- 17 variables

02

Geolocalización de Regiones

- Funcion de Geolocator
- Coordenadas regiones
- 953 regiones

03

Condiciones meteorológicas

- Meteostat API
- Rango 2015 - 2024
- Por estaciones del año

Variables Target

01

Volumen de Alcohol

Afecta el sabor y la sensación de calor en la garganta. Los vinos con alto contenido de alcohol se sienten más cálidos y pesados, mientras que los vinos con menor contenido son más ligeros y frescos.

- Grados de Alcohol

02

Cuerpo del vino

Peso y la densidad del vino en la boca. Los vinos ligeros son delicados y suaves, mientras que los vinos de cuerpo completo son robustos y densos. Influye en la textura y la sensación general del vino.

- High bodied | Medium bodied | Low bodied

03

Acidez del vino

Sensación de frescura y vivacidad en la boca, similar a morder una manzana verde. La acidez alta da una sensación crujiente y refrescante, equilibrando los sabores del vino.

- High | Medium | Low

Variables climatológicas

01

Temperatura media

tagv

02

Horas de sol

tsun

03

Precipitaciones

prcp

04

Nieve

snow

05

Presión atmosférica

pres

XXXX

1

**¿Cuáles son las correlaciones entre
los atributos del vino y las
condiciones meteorológicas?**

Volumen de Alcohol

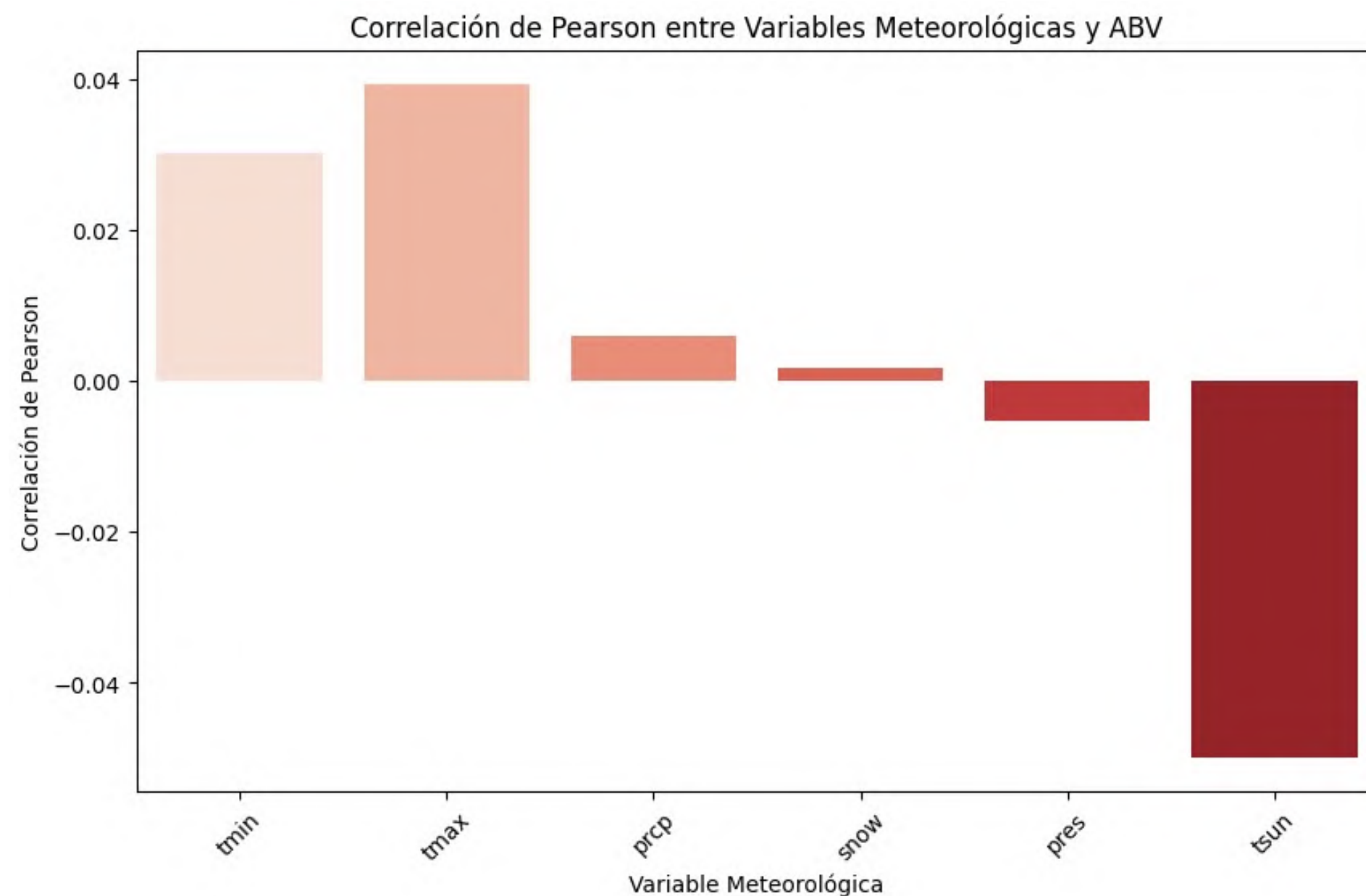
Cuerpo del Vino

Acidez

Correlación de Pearson

Variable numérica

Volumen de Alcohol



- **No hay una influencia significativa** en el contenido de **alcohol** de los vinos.
- Esto podría deberse a que el volumen de Alcohol está más influenciado por **factores controlables durante el proceso de fermentación y vinificación**, como la selección de levaduras y la duración de la fermentación, en lugar de las condiciones climáticas.

Volumen de Alcohol

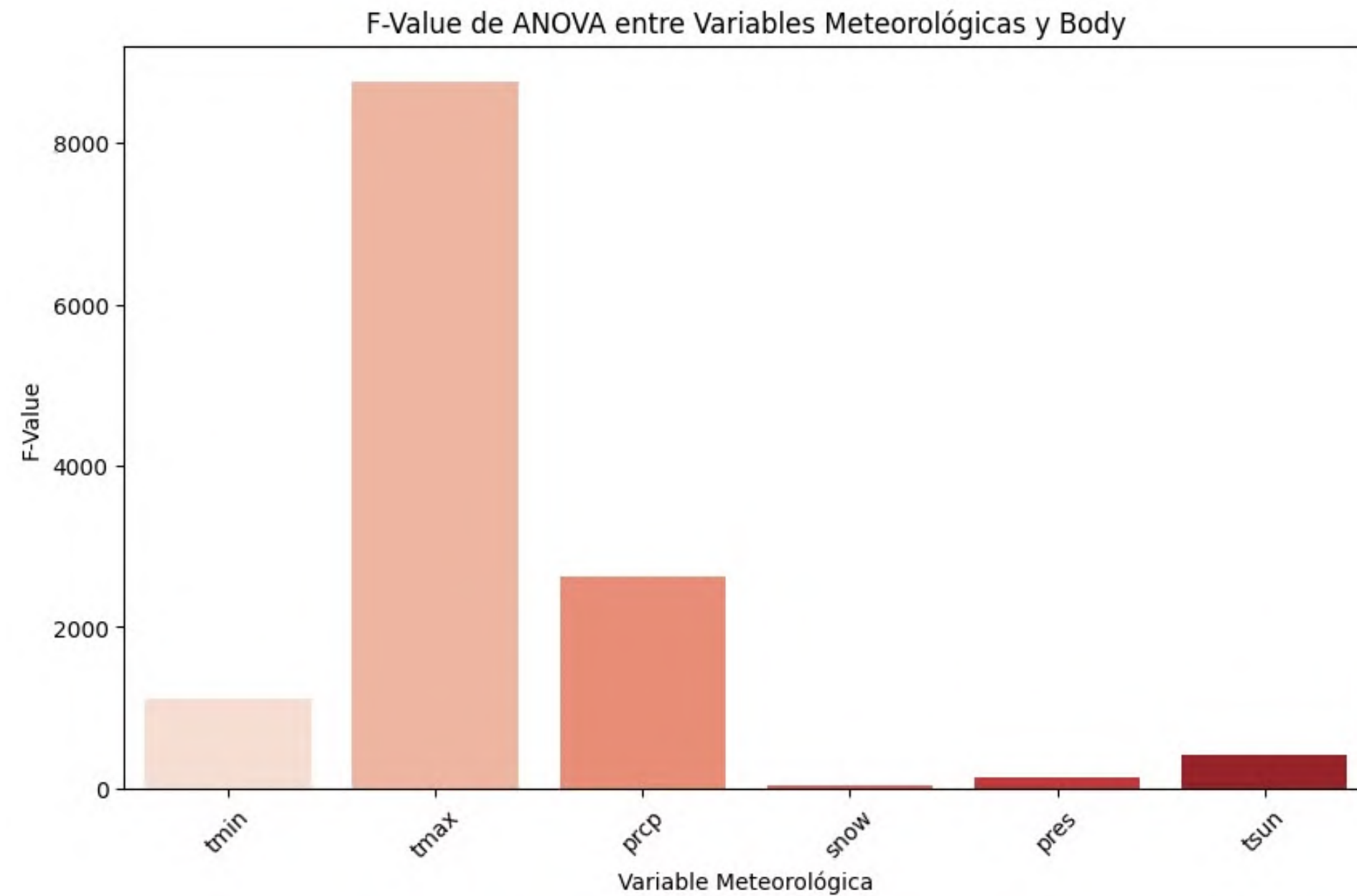
Cuerpo del Vino

Acidez

ANOVA

Variable categórica

Cuerpo



- La **temperatura máxima y la precipitación** tienen un impacto significativo en el cuerpo del vino.
- Esto podría indicar que en climas más cálidos y en regiones con mayor precipitación, **los vinos tienden a ser más robustos y con mayor cuerpo.**

Volumen de Alcohol

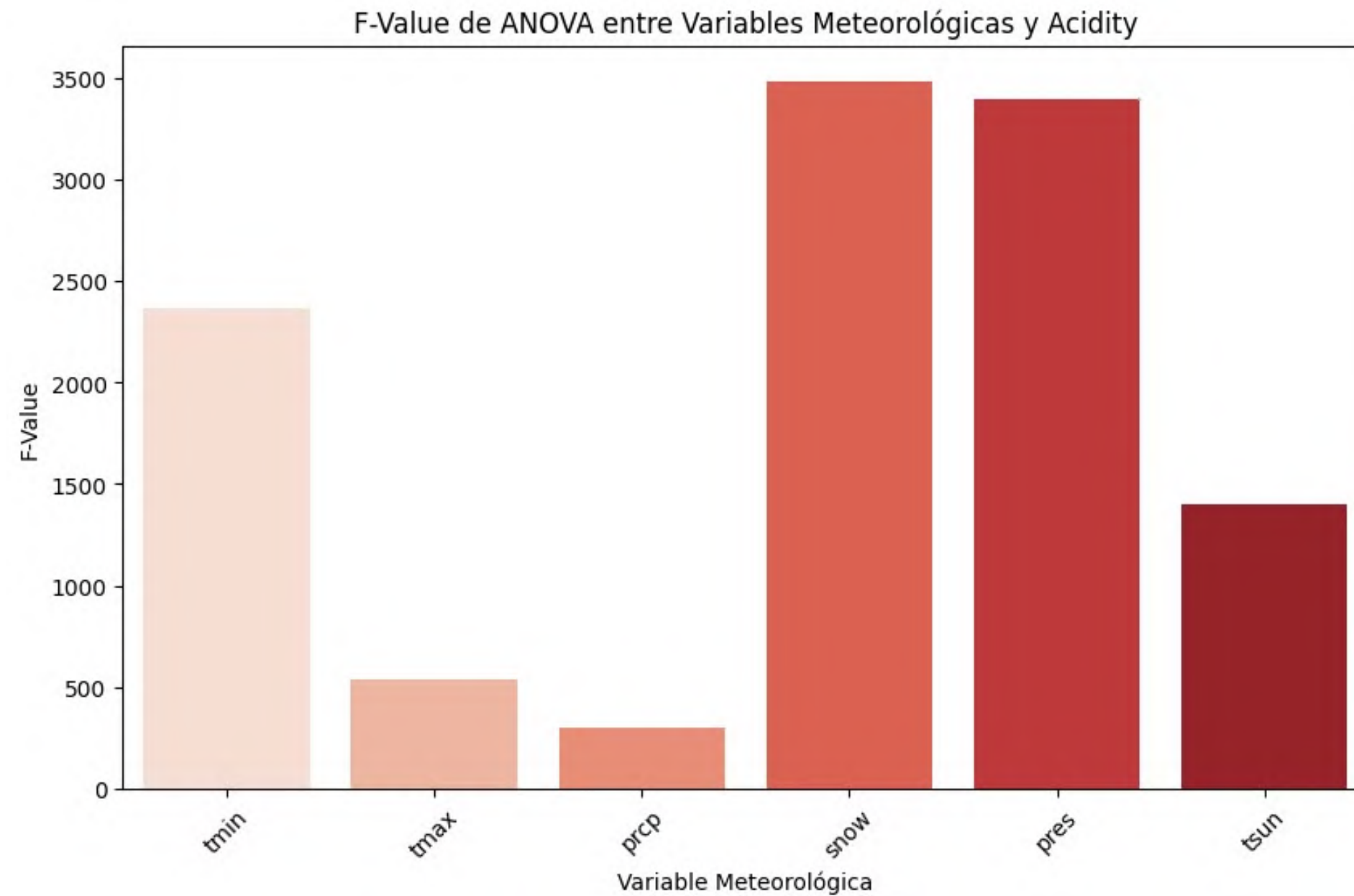
Cuerpo del Vino

Acidéz

ANOVA

Variable categórica

Acidéz



- La acidez del vino está significativamente influenciada por las condiciones climáticas, especialmente por la cantidad de **nieve y la presión atmosférica**.
- Esto sugiere que en regiones con más nieve y mayor presión atmosférica, **los vinos tienden a ser más ácidos**.
- Este efecto podría estar relacionado con cómo estas condiciones afectan el crecimiento y la maduración de las uvas, así como la retención de ácido en las uvas.

XXXX

3

**¿Se puede predecir los atributos
del vino en base a las
condiciones climatológicas?**

Volumen de Alcohol

Cuerpo del Vino

Acidez

Primeros modelos

Volumen de Alcohol

Afecta el sabor y la sensación de calor en la garganta. Los vinos con alto contenido de alcohol se sienten más cálidos y pesados, mientras que los vinos con menor contenido son más ligeros y frescos.

- Grados de Alcohol

- Primeros resultados para predecir el Volumen de Alcohol

Linear Regression

- Mean Squared Error: 48.73
- R^2 : 0.27

Decision Tree Regressor

- Mean Squared Error: 2.35
- R^2 : 0.96

Random Forest Regressor

- Mean Squared Error: 0.75
- R^2 : 0.97

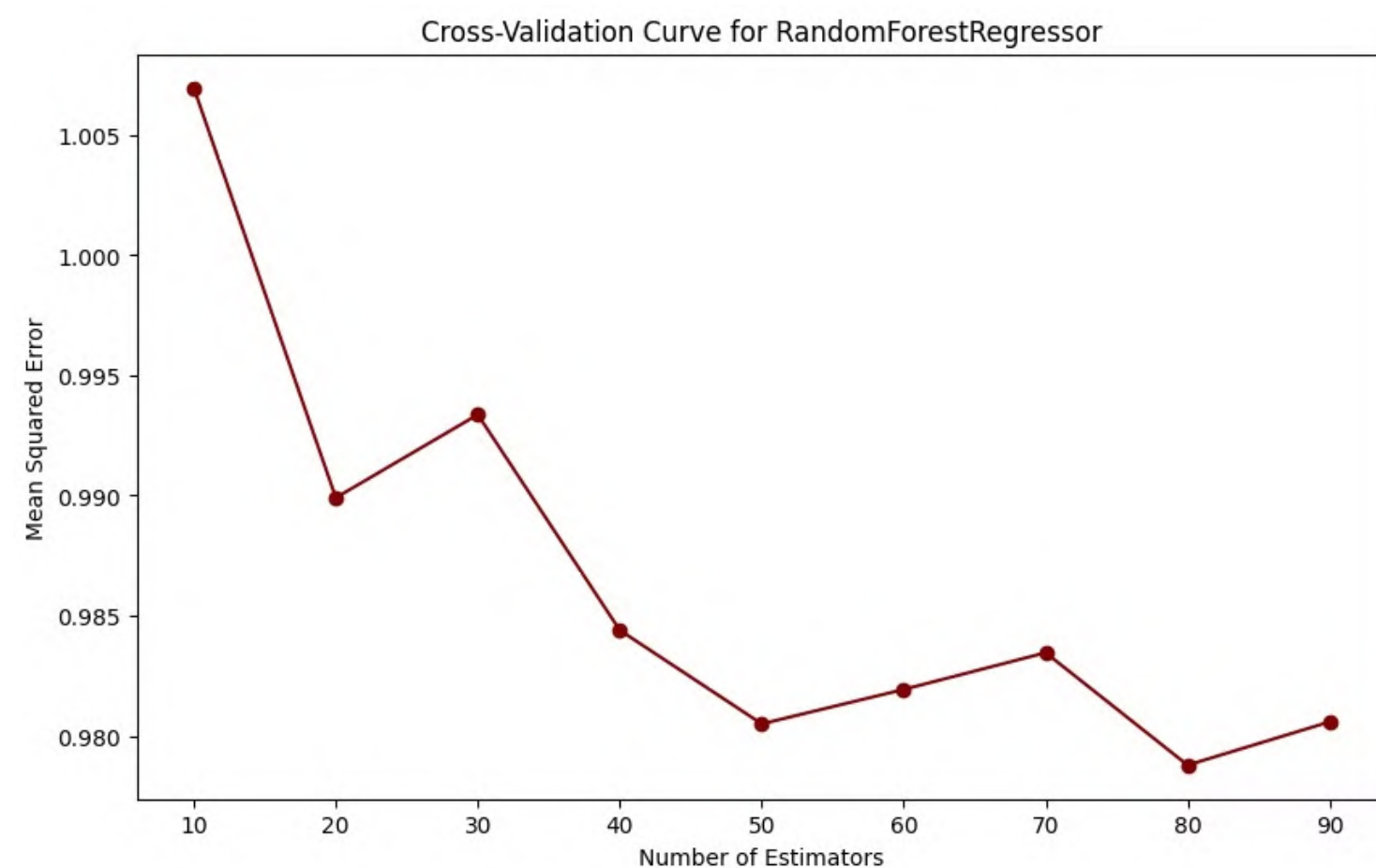
Volumen de Alcohol

Cuerpo del Vino

Acidéz

Mejoras del modelo

- **Cross validation** → identificar el número óptimo de estimadores que minimiza el error cuadrático medio (MSE)
optimal_estimators = 50
- **Correlaciones**: horas de exposición solar (tsun) vs temperatura media (tavg)
- Variable "season" numerica ordinal → pasarla a **dummie**



season_spring -

season_summer -

season_winter -

tavg -

prcp -

snow -

pres -

tsun -

Latitude -

Longitude -

time -

season_spring -	1.00	-0.27	-0.39	0.29	-0.57	-0.30	0.25	0.72	-0.02	-0.02
season_summer -	-0.27	1.00	-0.27	0.78	0.23	-0.21	-0.13	0.43	0.07	0.08
season_winter -	-0.39	-0.27	1.00	-0.70	-0.19	0.36	0.59	-0.60	-0.02	-0.02
tavg -	0.29	0.78	-0.70	1.00	0.03	-0.47	-0.18	0.81	-0.03	-0.02
prcp -	-0.57	0.23	-0.19	0.03	1.00	0.08	-0.69	-0.39	0.36	0.13
snow -	-0.30	-0.21	0.36	-0.47	0.08	1.00	-0.03	-0.42	0.05	0.34
pres -	0.25	-0.13	0.59	-0.18	-0.69	-0.03	1.00	0.12	-0.37	-0.28
tsun -	0.72	0.43	-0.60	0.81	-0.39	-0.42	0.12	1.00	-0.07	0.01
Latitude -	-0.02	0.07	-0.02	-0.03	0.36	0.05	-0.37	-0.07	1.00	0.66
Longitude -	-0.02	0.08	-0.02	-0.02	0.13	0.34	-0.28	0.01	0.66	1.00

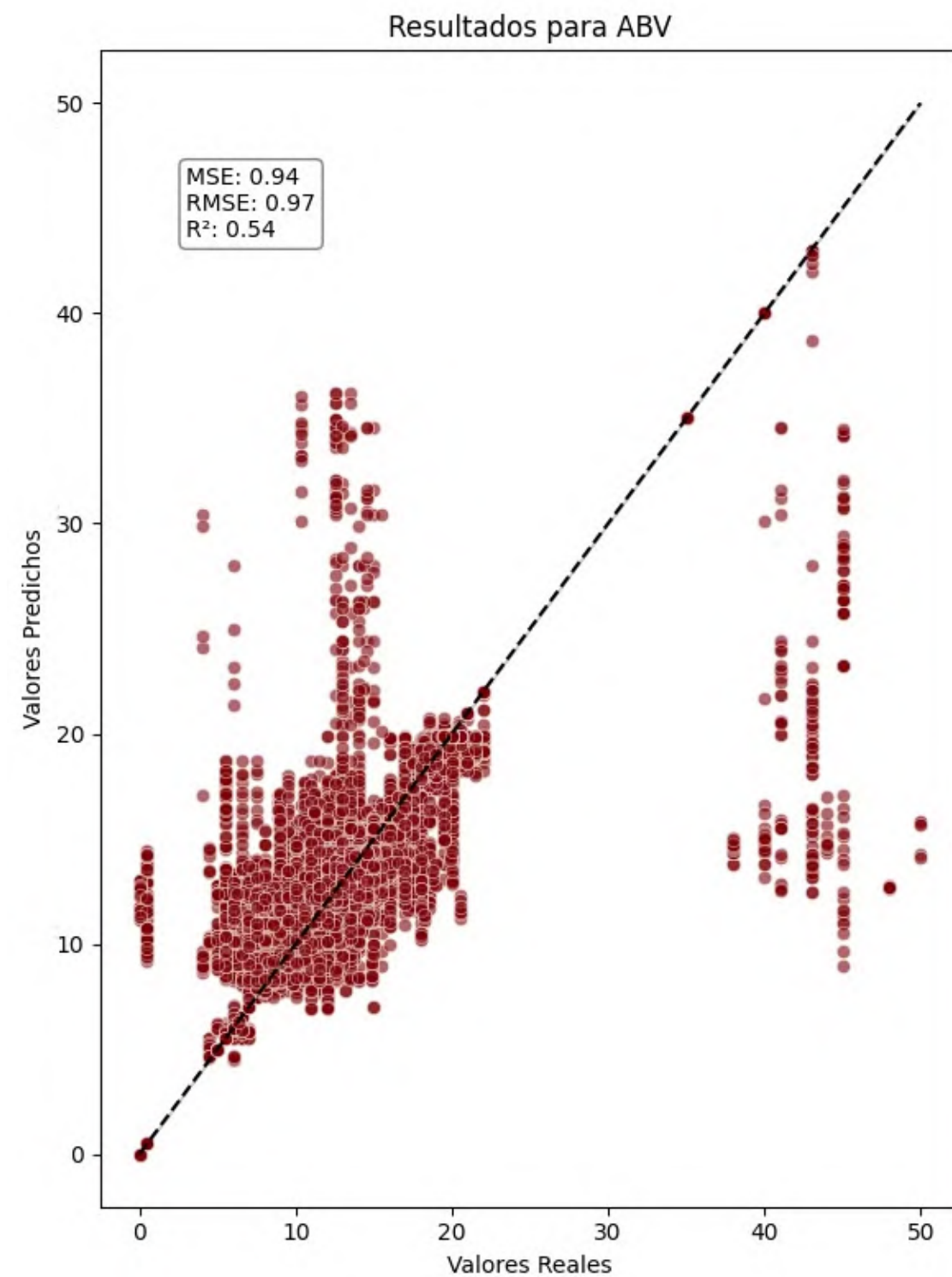
Volumen de Alcohol

Cuerpo del Vino

Acidez

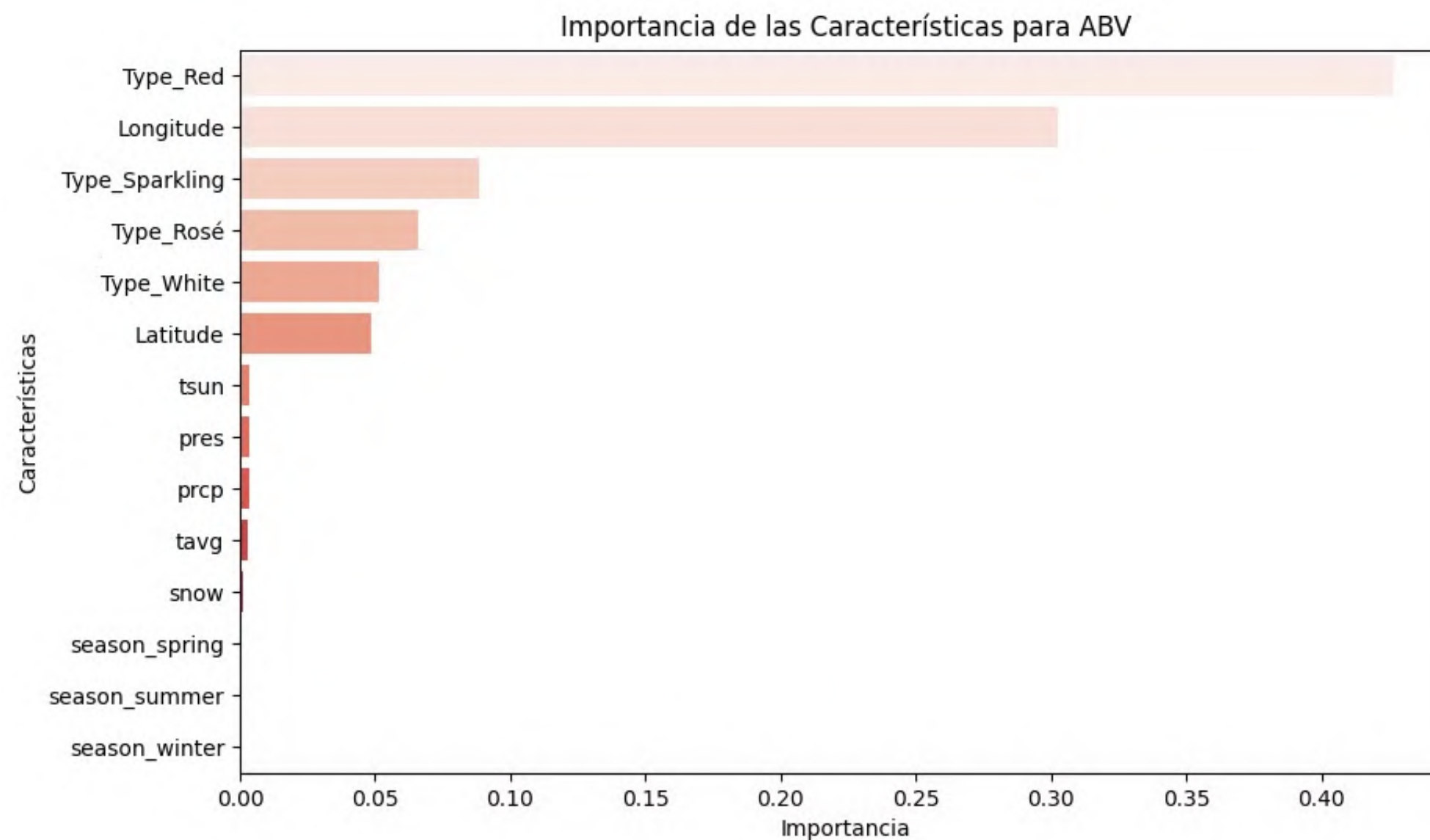
Modelo para predecir Volumen de Alcohol

RandomForestRegressor



Mean Squared Error (ABV): 0.94

R² (ABV): 0.54



Importancia de características: Para identificar cuáles son los factores climáticos y características del vino que tienen un mayor impacto en la predicción del ABV.

Volumen de Alcohol

Cuerpo del Vino

Acidéz

Modelo especializado por tipo de vino

RandomForestRegressor

[Tinto] Mean Squared Error (ABV): 0.49

[Tinto] R^2 (ABV): 0.33

- **Modelo Especializado**
- **Variabilidad Dentro de Subgrupos**
- **Overfitting a Subgrupos**

Volumen de Alcohol

Cuerpo del Vino

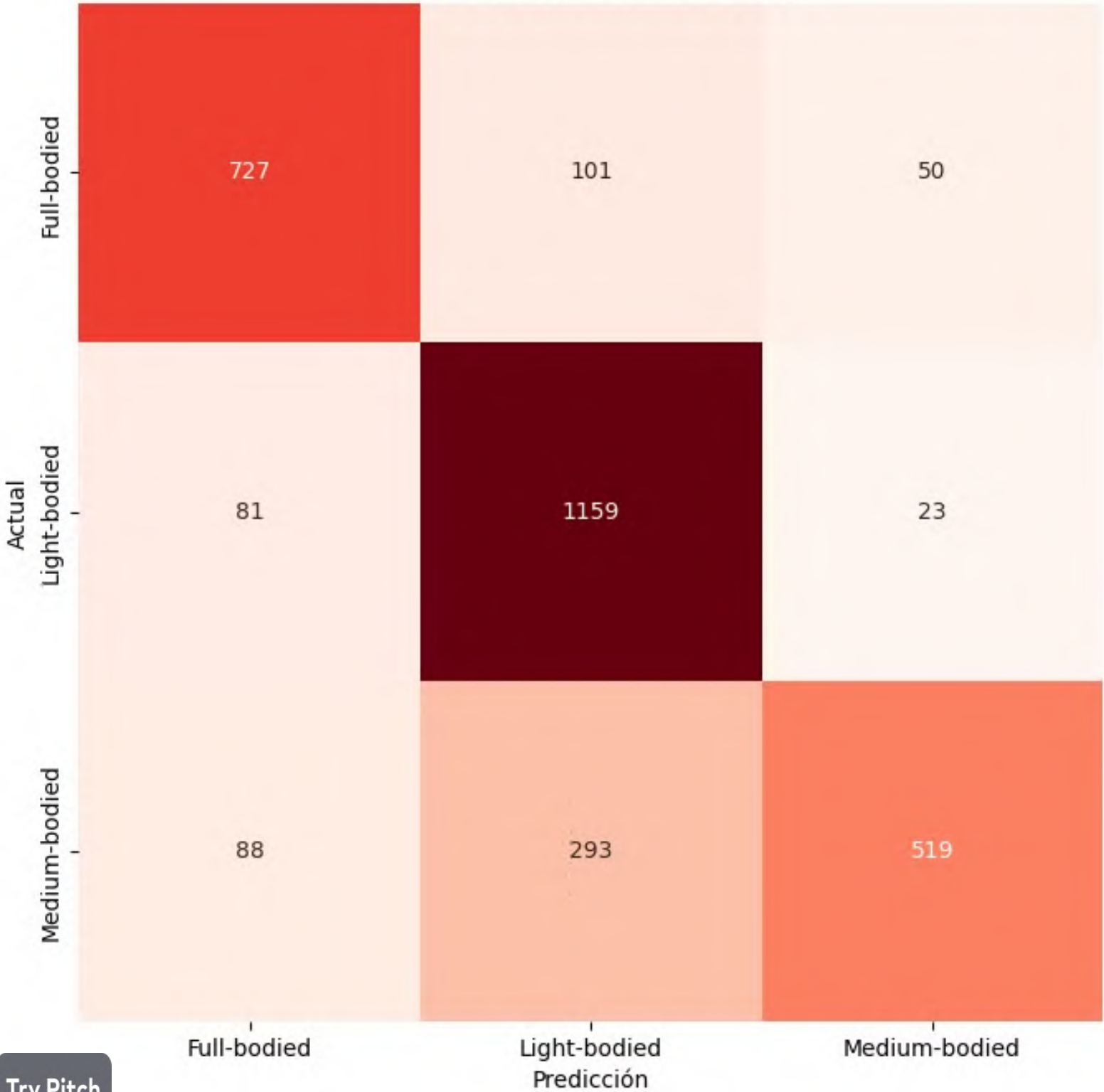
Acidéz

Modelo para predecir el Cuerpo y la Acidéz del vino

RandomForestClassifier

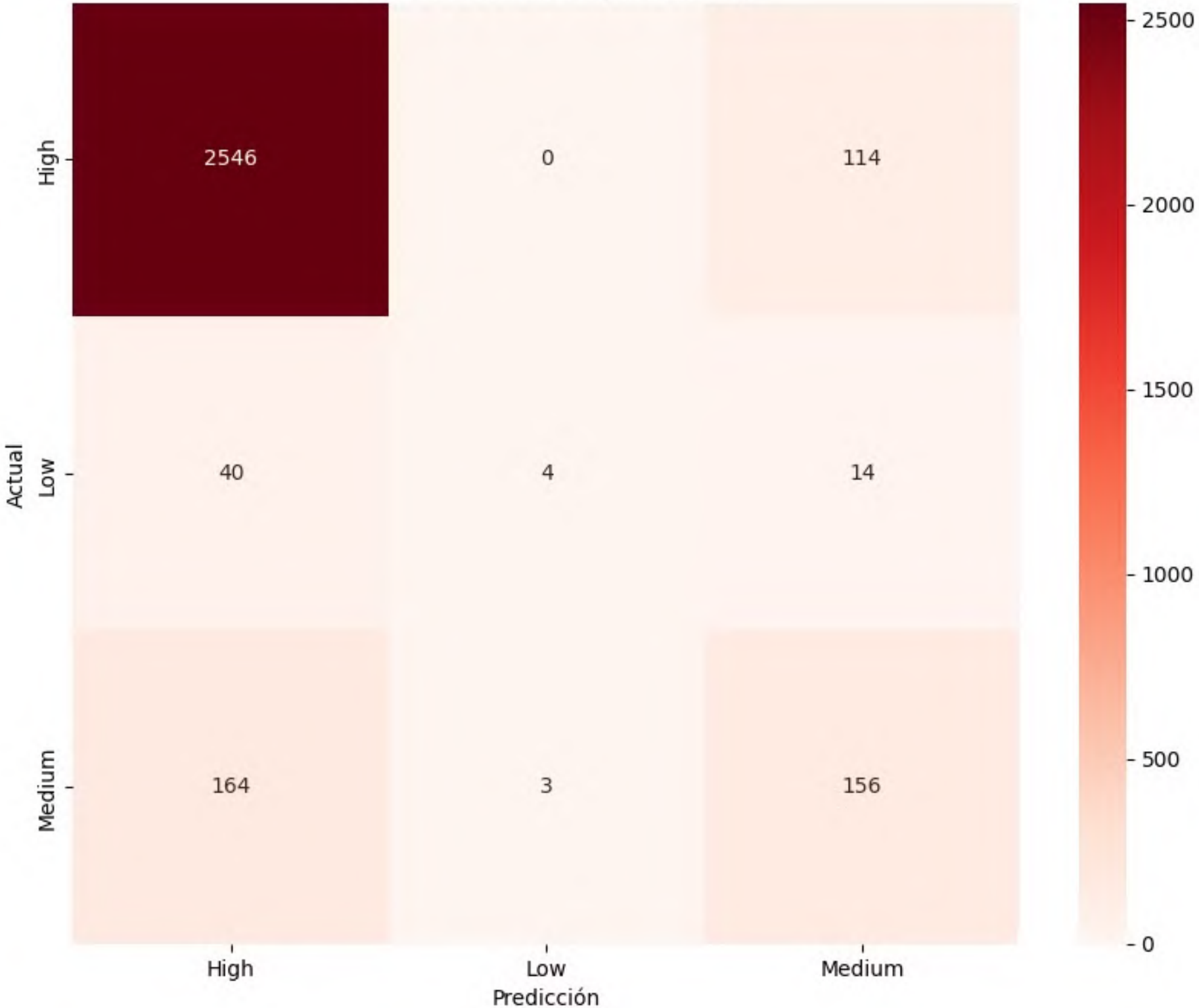
Cuerpo

Matriz de Confusión para Body



Acidéz

Matriz de Confusión para Acidity



Try Pitch

Accuracy: 0.79

Accuracy: 0.88

XXXX

4

¿Se puede **predecir el cambio
climático?**

Temperatura
media

Horas de sol

Presión
atmosférica

Precipitaciones

Nieve

Modelo para variables climáticas

Linear Regression

- Mean Squared Error: 43.87
- R^2 : 0.25

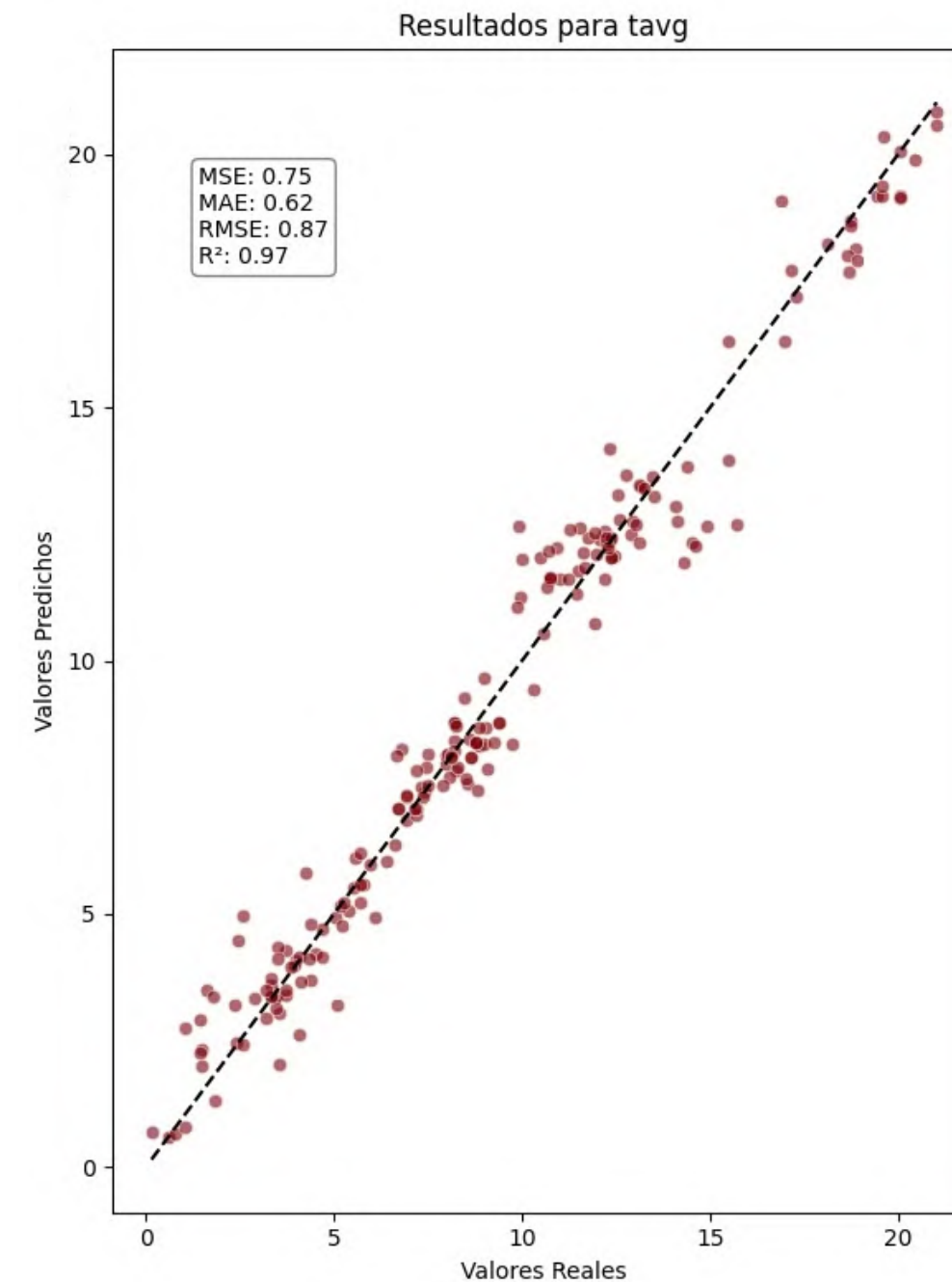
Decision Tree Regressor

- Mean Squared Error: 2.35
- R^2 : 0.95

Random Forest Regressor

- Mean Squared Error: 0.75
- R^2 : 0.97

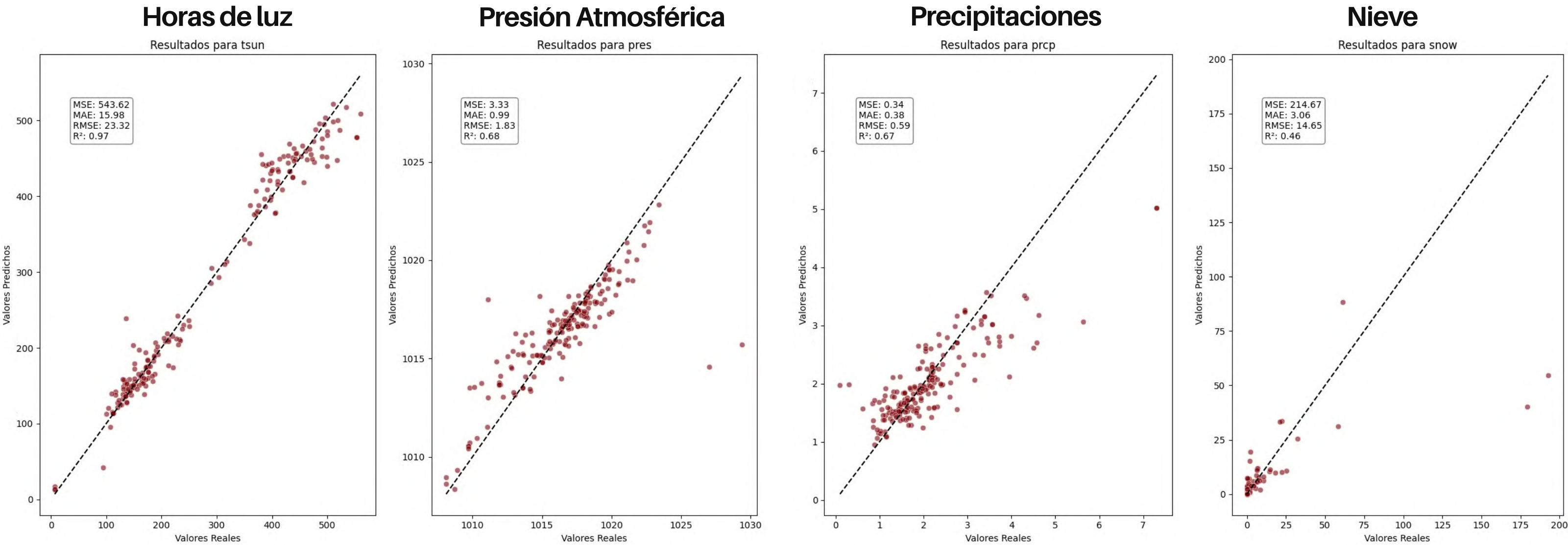
Temperatura media





Modelo

RandomForestRegressor



- Mean Squared Error: 543.67
 - R^2: 0.97
- Mean Squared Error: 3.33
 - R^2: 0.68
- Mean Squared Error: 0.34
 - R^2: 0.67
- Mean Squared Error: 214.67
 - R^2: 0.46

Conclusiones

01 **Cambio climático**

02 **Complejidad de la producción del vino**

03 **Limitación por el tipo de datos**



Want to make a presentation like this one?

Start with a fully customizable template, create a beautiful deck in minutes, then easily share it with anyone.

Create a presentation (It's free)