

UNIVERSIDAD  
**AUSTRAL**



**#30** INGENIERÍA

---

## ***Predicción de viajes en bicicleta en la ciudad de Rosario***

**Análisis de Series Temporales  
Maestría en Ciencia de Datos  
Universidad Austral**

Trabajo Práctico N°2

Autores:

Bermejo, Joaquín

Scarafia, Franco

Seward, Gerard

Año 2025

## Resumen Ejecutivo

Se propone en este trabajo predecir el número de viajes diarios (con el sistema público de bicicletas de la Municipalidad de Rosario) a tres estaciones cercanas a facultades de la UNR: Ciencias Económicas, Medicina e Ingeniería. En esta oportunidad se prueban modelos de *machine learning* para la predicción de series de tiempo, comparándolos entre sí y con los modelos tradicionales empleados en el estudio anterior.

## Índice de contenidos

<b>Resumen Ejecutivo</b>	<b>2</b>
<b>Índice de contenidos</b>	<b>3</b>
<b>Introducción</b>	<b>4</b>
<b>Marco Teórico</b>	<b>5</b>
<b>Análisis de resultados</b>	<b>7</b>
<b>Conclusiones</b>	<b>10</b>
<b>Referencias bibliográficas</b>	<b>11</b>

## Introducción

Desde 2015, el sistema público de bicicletas “Mi Bici Tu Bici” de la Municipalidad de Rosario ofrece un medio de transporte público alternativo dentro de la ciudad. El sistema funciona las 24 horas, los 365 días del año. A la actualidad, cuenta con más de cien mil usuarios y más de siete millones de viajes registrados (Mi Bici Tu Bici, 2024).

Con el objetivo de maximizar la accesibilidad al servicio, la empresa Movi realiza diariamente operaciones de balanceo, las cuales buscan que la mayor cantidad posible de estaciones y por la mayor cantidad posible de tiempo tengan bicicletas disponibles para ser retiradas por los usuarios y puntos de anclaje libres para que los mismos puedan devolver sus bicicletas al finalizar su viaje. Lógicamente, para realizar esta tarea es necesario conocer los comportamientos migratorios de las bicicletas a través de las distintas estaciones de la ciudad.

Se tiene un conjunto de datos con el número de viajes diarios realizados a cada estación, a partir de 2021. Específicamente, las variables registradas son:

- **fecha:** fecha en formato yyyy-mm-dd. Para cada estación de bicicletas, se tiene una fila por cada día desde 2021-01-01 hasta 2024-04-30.
- **destino:** la estación de bicicletas donde finalizaron los viajes registrados (ejemplo: “Plaza San Martín”).
- **viajes:** el número de viajes diarios que finalizaron en cada estación.

Con el fin de prevenir la saturación o escasez de bicicletas, se propone en este trabajo considerar tres estaciones y utilizar estos datos históricos para predecir el número de viajes diarios futuros a cada una de ellas. Las tres estaciones incluidas son:

- Facultad de Ciencias Económicas
- Facultad de Medicina
- Facultad de Ingeniería

El siguiente trabajo se encuentra dividido en tres secciones. La primera sección incluye una descripción de las series, la identificación de múltiples modelos candidatos para cada una, y una evaluación de su capacidad predictiva. La segunda sección se enfoca en la selección de uno de los modelos y los resultados obtenidos bajo éste. La tercera sección presenta las conclusiones obtenidas y las limitaciones de este trabajo, las cuales a su vez podrían motivar una investigación futura.

## Marco Teórico

Las Figuras 1, 2 y 3 muestran el comportamiento de las series escogidas. En todas ellas puede observarse una estacionalidad semestral: los meses de enero y julio registran un decrecimiento en el número de viajes diarios, quizás a causa de los recesos de verano e invierno en la universidad. A partir de 2023, el número medio de viajes diarios parece decrecer en las tres series.

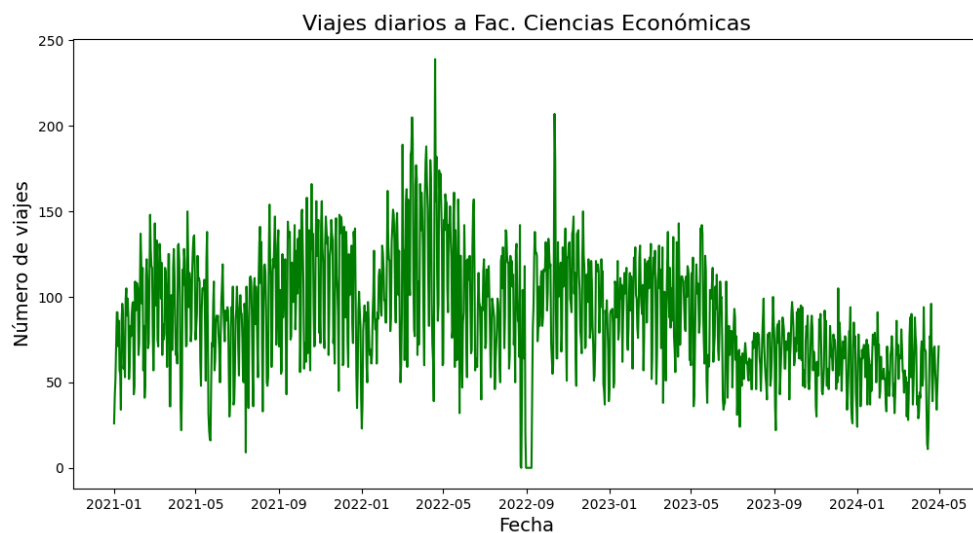


Figura 1. Número de viajes en bicicleta diarios a la Facultad de Ciencias Económicas

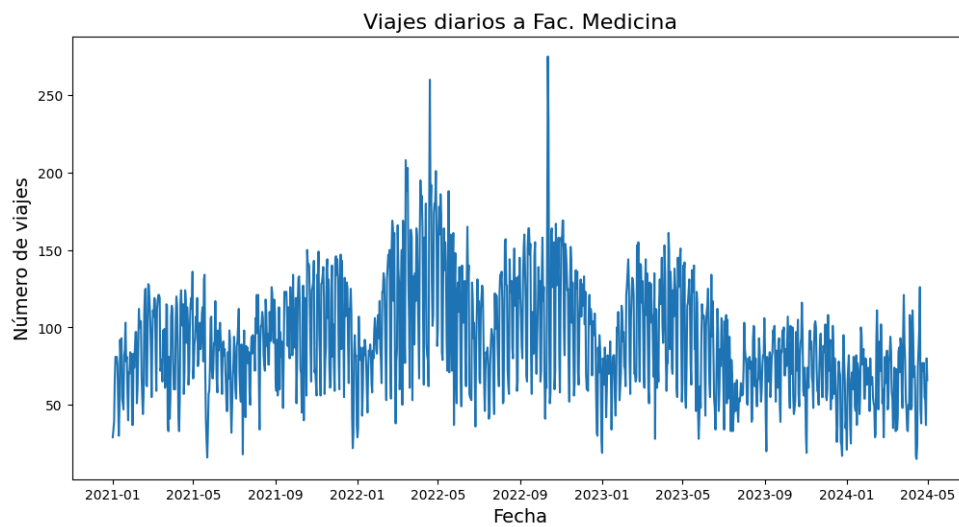


Figura 2. Número de viajes en bicicleta diarios a la Facultad de Medicina

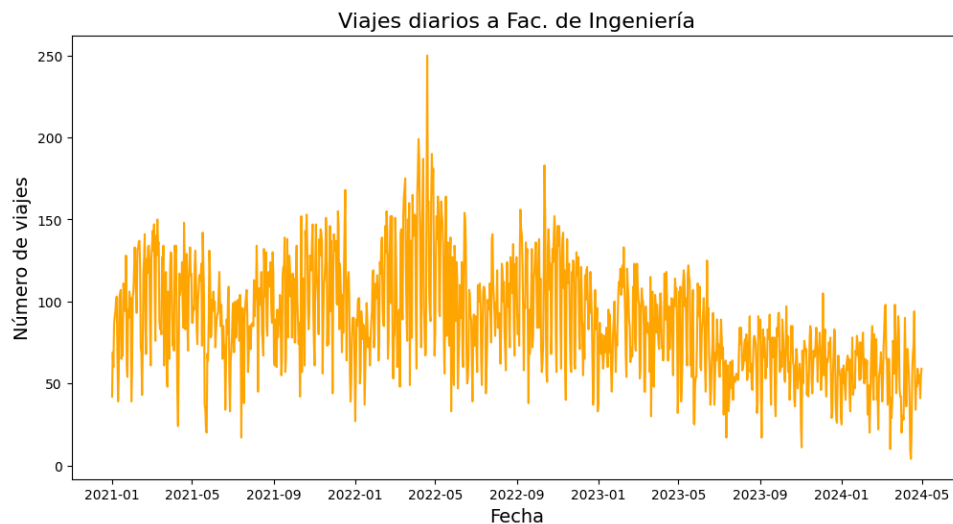


Figura 3. Número de viajes en bicicleta diarios a la Facultad de Ingeniería

En todas las estaciones se observan picos de actividad en tres días específicos:

- 19 de abril de 2022: viajes gratuitos en celebración del Día de la Bicicleta.
- 12 y 13 de octubre de 2022: paro de colectivos en Rosario.

Adicionalmente, para la Facultad de Ciencias Económicas se observa un número nulo de viajes durante un total de diez días, en agosto y septiembre de 2022. Esto puede deberse a un malfuncionamiento de la estación o a una pérdida de datos.

Las series no parecen ser estacionarias, si bien tampoco muestran una tendencia definida (como ya se mencionó, podría haber una leve tendencia negativa a partir de 2023).

## Análisis de resultados

Se ajustaron los siguientes modelos a cada una de las series presentadas:

- Prophet (modelo de Meta para el pronóstico de series de tiempo).
- Batería de modelos de AutoGluon (Chronos, DeepAR, TiDE, entre otros).

El modelo final es seleccionado en base a la bondad de las predicciones obtenidas y la relación costo computacional-mejora del modelo. Para llevar a cabo esta evaluación de la capacidad predictiva, se considera la siguiente partición del conjunto de datos:

- Entrenamiento: 01/01/2021 hasta 31/12/2023 (1095 días).
- Evaluación: 01/01/2024 hasta 30/04/2024 (121 días).

Para determinar si un modelo es bueno, se comparan los resultados contra los mejores modelos ajustados en el estudio anterior (SARIMA).

Una de las ventajas del modelo Prophet es la posibilidad de descomponer el comportamiento de una serie de tiempo en múltiples estacionalidades con diferentes períodos. La Figura 1 muestra la descomposición de la serie “Fac. Cs. Económicas”, si bien las tres series son muy similares y por lo tanto las siguientes observaciones aplican a todas ellas.

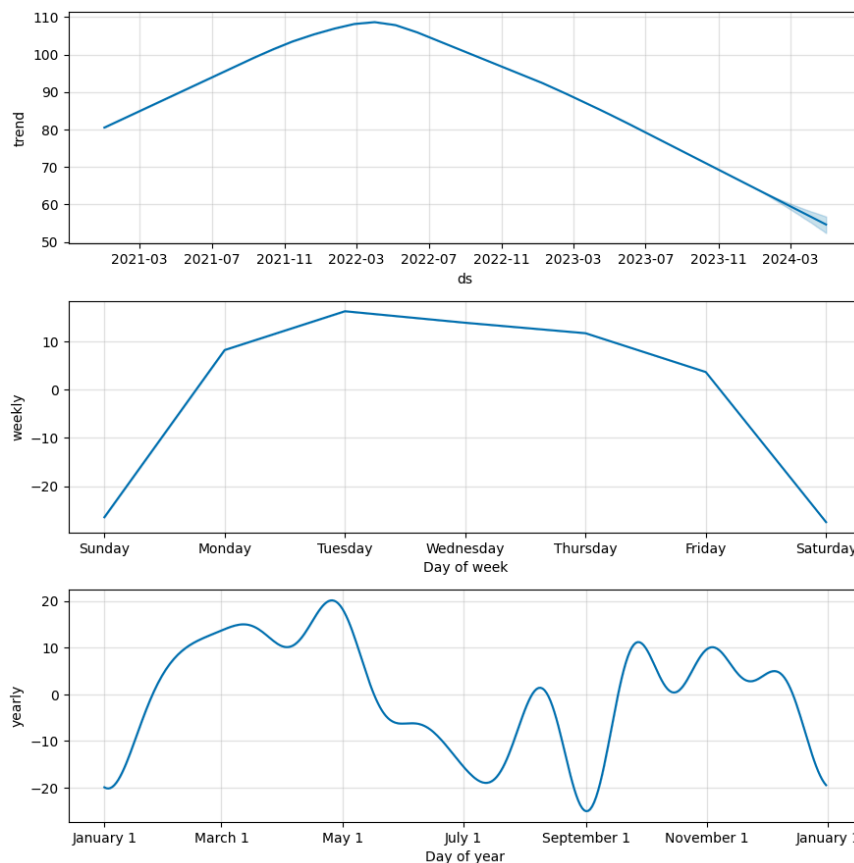


Figura 1. Descomposición de “Fac. Cs. Económicas” en tendencia y dos estacionalidades.

En primer lugar, se puede observar un cambio de tendencia durante el segundo trimestre de 2022. Antes de este período la serie mostraba una tendencia creciente, pero después de este período la tendencia se vuelve decreciente. El primer comportamiento podría estar explicado por la pandemia de COVID-19, la cual llevó a una cuarentena estricta que restringió la movilidad de los ciudadanos durante los años 2020 y 2021. Sobre la tendencia decreciente no se ha podido establecer un motivo. Sería conveniente aplicar esta descomposición a todas las estaciones de bicicletas en conjunto, para ver si se presenta sólo en las estaciones cercanas a una facultad o si fue un fenómeno común.

La tendencia semanal muestra un decrecimiento en el número de viajes diarios durante los fines de semana. Esto es de esperarse, visto que el número de personas que asiste a las facultades en estos días se ve reducido considerablemente.

Por último, la estacionalidad anual permite ver una reducción en el número de viajes diarios durante principios y mediados de año, lo cual nuevamente concuerda con los períodos de baja actividad en las facultades.

Luego de este análisis exploratorio, se procedió a ajustar los distintos modelos a los tres conjuntos de datos. La Tabla 1 muestra los valores RMSE de los modelos para los conjuntos de evaluación de cada una de las series bajo estudio. También se incluye la métrica agregada sobre las tres series, lo cual se calcula mediante la raíz cuadrada del promedio de errores cuadráticos medios. La tabla se encuentra ordenada en forma decreciente de desempeño general.

Los modelos SARIMA ocupan el tercer puesto en la lista. Tan solo dos modelos logran un mejor desempeño, aunque no para las tres estaciones de bicicletas. El mejor de ellos, sin embargo, supone una mejora de apenas un 2,5% sobre los modelos SARIMA.

Nótese además que, de los 13 modelos de machine learning evaluados en este ensayo, cinco de ellos tienen un desempeño inferior al pronóstico naïve estacional. Estos modelos no merecen ser considerados para predecir los viajes diarios en bicicleta, puesto que son relativamente complejos (es decir, entrenarlos y ponerlos en producción conllevaría un cierto costo computacional) y aún así no logran superar un modelo extremadamente sencillo.



Modelo \ Serie	General	Fac. Cs. Económicas	Fac. Medicina	Fac. de Ingeniería
ChronosFineTuned[bolt_small]	17.43 (-2.5%)	15.73 (-6.5%)	19.46 (+2.4%)	16.89 (-4.8%)
AutoETS	17.50 (-2.1%)	15.45 (-8.1%)	19.94 (+4.9%)	16.82 (-5.2%)
<b>Modelos SARIMA</b>	<b>17.88</b>	<b>16.82</b>	<b>19.00</b>	<b>17.74</b>
DynamicOptimizedTheta	18.49 (+3.4%)	18.58 (+10%)	19.81 (+4.3%)	16.96 (-4.4%)
ChronosZeroShot[bolt_base]	18.96 (+6.0%)	18.13 (+7.8%)	20.42 (+7.5%)	18.25 (+2.9%)
PatchTST	19.34 (+8.2%)	18.73 (+11%)	18.85 (-0.8%)	20.41 (+15%)
TiDE	20.59 (+15%)	20.08 (+19%)	21.16 (+11%)	20.52 (+16%)
Prophet	21.76 (+22%)	21.95 (+30%)	22.66 (+19%)	20.63 (+16%)
WeightedEnsemble	23.02 (+29%)	24.31 (+45%)	23.09 (+22%)	21.60 (+22%)
SeasonalNaive	23.80 (+33%)	21.05 (+25%)	28.6 (+51%)	20.93 (+18%)
TemporalFusionTransformer	26.29 (+47%)	26.95 (+60%)	26.57 (+40%)	25.32 (+43%)
RecursiveTabular	29.38 (+64%)	31.52 (+87%)	31.29 (+65%)	24.84 (+40%)
NPTS	35.88 (+101%)	37.07 (+120%)	33.22 (+75%)	37.19 (+110%)
DeepAR	37.39 (+109%)	38.87 (+131%)	36.98 (+95%)	36.27 (+104%)
DirectTabular	38.66 (+116%)	40.26 (+139%)	32.73 (+72%)	42.34 (+139%)

Tabla 1. Evaluación de cada modelo (métrica: RMSE).

## Conclusiones

En esta segunda iteración sobre la evaluación de modelos de series de tiempo, se han puesto a prueba más herramientas de forecasting para predecir el número diario de viajes en bicicleta utilizando el sistema público de la Municipalidad de Rosario.

Al igual que en el estudio anterior, se ha apreciado que la complejidad de los modelos no necesariamente se traduce en mejores resultados. Los modelos SARIMA, sobre los cuales se hizo foco en el primer ensayo, se encuentran entre los mejores modelos evaluados. Debe recordarse, sin embargo, que éstos no verificaban todos los supuestos necesarios.

La aplicación de modelos sofisticados, sin embargo, aporta valor no sólo a la predicción de valores futuros, sino al análisis exploratorio y la caracterización de una serie. Como se ha visto, el uso del modelo Prophet para descomponer las series en sus múltiples estacionalidades resultó de utilidad para detectar patrones migratorios en los usuarios, algunos de los cuales son esperados y otros que requieren una mayor profundización para hallar la causa raíz.

Este segundo ensayo refuerza la conclusión obtenida en el estudio anterior: que el número de viajes diarios es una variable afectada por múltiples fenómenos, los cuales no están presentes en el conjunto de datos. En consecuencia, se especula que una mejora sustancial en la predicción de viajes diarios no depende de la elección del modelo "correcto", sino de una recolección de datos estratégica y un *feature engineering* hecho a medida para el problema en cuestión: el balanceo de las estaciones de bicicletas.

## Referencias bibliográficas

- Del Rosso, R., y Drago, B. (2024) *Análisis de Series Temporales*. Apuntes de cátedra.
- Mi Bici Tu Bici. (2024). *El sistema*. Mi Bici Tu Bici.  
<https://www.mibicitubici.gob.ar/#!/app/home>