

Universidad Tecnológica de Panamá  
Facultad de Ingeniería de Sistemas Computacionales

**Proyecto Final Modelos Predictivos**

**Docente:**

**Dr. Juan Castillo**

**Estudiante**

Rodrigo Yánguez

Maestría en Analítica de Datos

2024

## Contenido

Introducción .....	3
Problemática .....	3
Metodología .....	4
Análisis de monto y cantidad de beneficiarios por año .....	5
Predicción de monto Asignado a beneficiarios de la Universidad Tecnológica de Panamá ..	7
Promedio Móvil.....	8
Suavización exponencial.....	9
Modelo de Holt .....	10
Modelo ARIMA .....	11
Predicción de Cantidad de Beneficiarios de la Universidad Tecnológica de Panamá .....	12
Promedio Móvil.....	12
Suavización Exponencial.....	14
Modelo Holt.....	14
Modelo ARIMA .....	15
Resultados .....	16
Hallazgos.....	19
Conclusiones.....	20
Bibliografía .....	20

## Introducción

En el desarrollo de la asignatura Modelos Predictivos, se estuvo revisando múltiples casos de ejemplo, muchos basados en información o casos reales, pero al fin al cabo ejemplos que los libros de texto utilizan para el desarrollo de las habilidades de los estudiantes. Sin embargo, es realmente desafiante cuando se lleva este conocimiento a información real de situaciones puntuales en diversos ámbitos del conocimiento.

Los modelos predictivos ayudan a valorar los resultados futuros basados en el histórico de datos. Es importante resaltar que dichos datos deben tener un volumen importante para que el modelo se entrene correctamente.

A lo largo de este trabajo se estará presentando la problemática a abordar, las razones por la cual se escogió el análisis, los modelos utilizados y la interpretación de los resultados. Además, se ha añadido una sección de hallazgos para mostrar algunas situaciones que llamaron la atención durante el análisis y que deben ser tomadas en cuenta a la hora de presentar este trabajo.

Por último, se presentan las conclusiones acerca del trabajo realizado, resaltando las principales enseñanzas para nuestro entendimiento.

## Problemática

El Instituto para la Formación y Aprovechamiento de Recursos Humanos (IFARHU) es la institución del estado panameño dedicada al desarrollo de la formación y aprovechamiento del capital humano, mediante programas y proyectos destinados a la comunidad en general, otorgando becas y créditos educativos para contribuir al mejoramiento de la calidad de vida de la población nacional (IFARHU, s.f.).

Anualmente se destinan millones de dolares en presupuestos a esta institución para el otorgamiento de becas, préstamos e incentivos a estudiantes panameños de todos los niveles de enseñanza (primaria, secundaria, superior) para estudios en instituciones educativas nacionales e internacionales. También se financian capacitaciones para funcionarios públicos del gobierno panameño a nivel de cursos, diplomados y especialidades.

Sin embargo, en los últimos años ha habido cierta polémica en la forma en que se otorgan estos beneficios, principalmente en el programa de Auxilios Económicos.

Es por ello que el IFARHU ha publicado un dataset con la información de todos los beneficios concedidos por la institución en el período 2014-2024. Dicho dataset es una fuente valiosa de información para entender no solamente los montos otorgados, sino también los intereses de los estudiantes panameños sobre temas como universidades, carreras y países.

Luego de hacer un estudio previo de los datos, y con la orientación del profesor, se decidió analizar los datos referentes a los beneficios otorgados a estudiantes de Universidad Tecnológica de Panamá para poder predecir el monto que debe ser otorgado durante el 2024 y la cantidad de estudiantes beneficiados por los mismos.

Asimismo, se analizó la tendencia entre el monto de dinero asignado vs cantidad de beneficiarios, a lo largo del periodo 2015-2023 con el fin de generar los indicadores clave como por ejemplo el promedio del monto otorgado a cada estudiante, que puedan ayudar a analizar el comportamiento histórico de estos préstamos.

## Metodología

En primer lugar, se descargó el dataset del sitio web del IFARHU en formato xlsx. Este archivo se abrió en Excel, donde se llevaron a cabo las acciones de filtrado y limpieza de datos. Posteriormente, utilizando tablas dinámicas, se agruparon y graficaron los datos para comprender su comportamiento. Este análisis permitió definir los objetivos del estudio. A continuación, se aplicaron los modelos predictivos de Promedio Móvil, Suavización Exponencial y Holt en Excel, y el método ARIMA en Python, utilizando Jupyter y Anaconda con la librería Pandas.

En la figura N° 1 se puede observar parte del dataset original publicado por el IFARHU

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Beneficiario	Concurso	Centro Educativo	País	Fecha de Inicio	Fecha de Finalización	Monto Otorgado	Monto Devuelto	Año	Universidad										
1	AEC CONCURSO NACIONAL	-NO LISTADA-	GENERAL CUL TURAL INSTITUTE	IRLANDA	02-2024	07-2024	25000	25000.00	2024	GENERAL CUL TURAL INSTITUTE									
2	CONVOCAATORIA PARA CONCURSO NACIONAL	MAESTRIA GESTION Y ADMON DEL CAPITAL HUMANO EN L	EUROPEAN OPEN BUSINESS SCHOOL (EOBS)	ESPAÑA	07-2022	07-2023	10400	10400.00	2022	EUROPEAN OPEN BUSINESS SCHOOL (EOBS)									
3	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA INGENIERIA AGRONOMICA	UNIV. ZAPORANO AGROPECUARIO	HONDURAS	01-2016	12-2019	68700	10000.00	2016	UNIV. ZAPORANO AGROPECUARIO									
5	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA INGENIERIA AGRONOMICA	ESC. AGRICOLA PANAMERICANA EL ZAPORANO	HONDURAS	01-2017	12-2018	36400	36400.00	2017	ESC. AGRICOLA PANAMERICANA EL ZAPORANO									
6	AEC CONCURSO NACIONAL	-NO LISTADA-	BAYSWATER COLLEGE - LONDRES	INGLATERRA	02-2022	10-2022	22209	22209.00	2022	BAYSWATER COLLEGE									
7	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA INGENIERIA AGRONOMICA	ESC. AGRICOLA PANAMERICANA EL ZAPORANO	HONDURAS	01-2017	12-2017	18200	16200.00	2017	ESC. AGRICOLA PANAMERICANA EL ZAPORANO									
8	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA INGENIERIA AGRONOMICA	UNIV. ZAPORANO AGROPECUARIO	HONDURAS	01-2016	12-2019	68700	72000.00	2016	UNIV. ZAPORANO AGROPECUARIO									
9	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA EN INGENIERIA AGRONOMO	ESC. AGRICOLA PANAMERICANA EL ZAPORANO	HONDURAS	01-2022	12-2023	38400	38400.00	2022	ESC. AGRICOLA PANAMERICANA EL ZAPORANO									
10	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA EN INGENIERIA AGRONOMO	UNIV. ZAPORANO AGROPECUARIO	HONDURAS	01-2022	12-2023	38400	38400.00	2022	UNIV. ZAPORANO AGROPECUARIO									
11	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA EN INGENIERIA AGRONOMO	ESC. AGRICOLA PANAMERICANA EL ZAPORANO	HONDURAS	01-2022	12-2023	38400	38400.00	2022	ESC. AGRICOLA PANAMERICANA EL ZAPORANO									
12	IFARHU - HEWLETT PACKARD	LICENCIATURA COMPORTAMIENTO ORGANIZAC Y DESAR	UNIV. INTERAMERICANA DE PANAMA - PANAMA	PANAMA	01-2014	09-2017	6691.6	6691.60	2014	UNIV. INTERAMERICANA DE PANAMA									
13	AEC CONCURSO NACIONAL	MAESTRIA ADMINISTRACION DE NEGOCIOS CON ENF EN TI	UNIV. INTERAMERICANA DE PANAMA - PANAMA	PANAMA	03-2019	08-2020	2900.4	2900.40	2019	UNIV. INTERAMERICANA DE PANAMA									
14	CONV DE COOP AGAD-BANCO NACIONAL	LICENCIATURA EN INGENIERIA AGRONOMICA ESP EN DE	ESC. AGRICOLA PANAMERICANA EL ZAPORANO	HONDURAS	01-2017	12-2018	36400	36400.00	2017	ESC. AGRICOLA PANAMERICANA EL ZAPORANO									
15	CONVOCAATORIA PARA CONCURSO NACIONAL	CURSO IDIOMA ALEMAN	UNIVERSIDAD DE PANAMA - CAMPUS	PANAMA	04-2018	01-2019	24303	24303.00	2018	UNIVERSIDAD DE PANAMA									
16	CONVOCAATORIA PARA CONCURSO NACIONAL	CURSO IDIOMA ALEMAN	UNIVERSIDAD DE PANAMA - CAMPUS	PANAMA	04-2018	01-2019	24303	24303.00	2018	UNIVERSIDAD DE PANAMA									
17	CONVOCAATORIA PARA CONCURSO NACIONAL	CURSO IDIOMA ALEMAN	UNIVERSIDAD DE PANAMA - CAMPUS	PANAMA	04-2018	01-2019	23953	23953.00	2018	UNIVERSIDAD DE PANAMA									
18	INTERES SOCIAL COMUNITARIO	LICENCIATURA EN INGENIERIA MECATRONICA	UNIVERSITY OF GLASGOW	SUECIA	09-2014	12-2014	20000	20000.00	2014	UNIVERSITY OF GLASGOW									
19	CONVOCAATORIA PARA CONCURSO NACIONAL	CURSO IDIOMA ALEMAN	UNIVERSIDAD DE PANAMA - CAMPUS	PANAMA	04-2018	01-2019	23953	23953.00	2018	UNIVERSIDAD DE PANAMA									
20	IFARHU - HEWLETT PACKARD	LICENCIATURA PRECULOSIA	UNIV. SANTA MARIA LA ANTIGUA - PANAMA	PANAMA	05-2014	04-2018	14910	14944.70	2014	UNIV. SANTA MARIA LA ANTIGUA									
21	INTERES SOCIAL COMUNITARIO	LICENCIATURA CONTABILIDAD	UNIV. ESPEC DEL CONTADOR PUBLICO AUTORIZADO	PANAMA	04-2014	12-2014	10000	10000.00	2014	UNIV. ESPEC DEL CONTADOR PUBLICO AUTORIZADO									
22	IFARHU - HEWLETT PACKARD	LICENCIATURA PUBLICIDAD Y MERCADO	UNIV. LATINA DE PANAMA - PANAMA	PANAMA	01-2014	04-2018	4677.94	5094.78	2014	UNIV. LATINA DE PANAMA									
23	IFARHU - HEWLETT PACKARD	LICENCIATURA EN INGENIERIA DE GAS	UNIV. LATINA DE PANAMA - PANAMA	PANAMA	01-2014	04-2018	6332.61	7136.90	2014	UNIV. LATINA DE PANAMA									
24	INTERES SOCIAL COMUNITARIO	MAE EN CIEN TIERRA Y DEL AMBIENTE ECOLOGIA ESP ECOTON	UNIV. LATINA DE PANAMA - PANAMA	PANAMA	05-2014	12-2014	7494.21	7494.21	2014	UNIV. LATINA DE PANAMA									
25	AUXILIO ECONOMICO	LICENCIATURA EN ENFERMERIA	BRENDA UNIVERSITY	ESTADOS UNIDOS	03-2022	12-2026	189560	117286.00	2022	BRENDA UNIVERSITY									
26	AUXILIO ECONOMICO COMPLEMENTARIO	LICENCIATURA EN DOCTOR EN MEDICINA	UNIVERSIDAD AUTONOMA DE CHIRIQUE - PANAMA	PANAMA	01-2021	06-2026	34385	15620.46	2021	UNIVERSIDAD AUTONOMA DE CHIRIQUE									
27	CONCURSO TRANSFORMANDO VIDAS	LICENCIATURA MERCADO Y PUBLICIDAD	WILKES UNIVERSITY	ESTADOS UNIDOS	01-2020	12-2024	192841.2	154113.96	2020	WILKES UNIVERSITY									
28	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	05-2021	07-2026	103742	84244.68	2021	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
29	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	03-2021	07-2026	103742	84844.70	2021	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
30	AUXILIO ECONOMICO	LICENCIATURA EN ENFERMERIA	UNIV. AMERICANA - PANAMA	PANAMA	03-2021	05-2025	21562.7	16846.70	2021	UNIV. AMERICANA									
31	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN ENFERMERIA	UNIV. INTERAMERICANA DE PANAMA - PANAMA	PANAMA	03-2021	03-2025	24436.7	13498.70	2021	UNIV. INTERAMERICANA DE PANAMA									
32	AUXILIO ECONOMICO COMPLEMENTARIO	LICENCIATURA EN DOCTOR EN MEDICINA	UNIVERSIDAD AUTONOMA DE CHIRIQUE - PANAMA	PANAMA	12-2020	12-2026	37374	21184.00	2020	UNIVERSIDAD AUTONOMA DE CHIRIQUE									
33	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	03-2021	07-2026	103742	84244.68	2021	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
34	AUXILIO ECONOMICO	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. INTERAMERICANA DE PANAMA - PANAMA	PANAMA	05-2022	04-2027	48000	15000.00	2022	UNIV. INTERAMERICANA DE PANAMA									
35	CONCURSO TRANSFORMANDO VIDAS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	09-2023	12-2029	105700.4	85550.40	2023	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
36	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	03-2021	07-2026	103742	81213.30	2021	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
37	AUXILIO ECONOMICO	LICENCIATURA DESARROLLO DE SOFTWARE	UNIV. TECNOLOGICA METROPOLITANA	CHILE	10-2020	12-2024	29000	14444.46	2020	UNIV. TECNOLOGICA METROPOLITANA									
38	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	05-2022	04-2029	103742	82247.38	2022	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
39	CONVOCAATORIA PARA OTROS CONVENIOS	LICENCIATURA EN DOCTOR EN MEDICINA	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	CUBA	05-2022	04-2029	103742	46971.02	2022	UNIV. DE CIENCIAS MEDICAS DE LA HABANA									
40	UNIV. DE CIENCIAS MEDICAS DE LA HABANA	LICENCIATURA EN INGENIERIA	UNIV. LATINA DE PANAMA - PANAMA	PANAMA	05-2020	10-2025	10000	7000.00	2020	UNIV. LATINA DE PANAMA									

Figura #1: Dataset del Ifarhu. Elaboración Propia.

Algunas de las acciones de limpieza de datos realizadas consistieron en:

- Verificación y eliminación de datos duplicados.
- Conversión de los valores de monto de tipo Texto a tipo Número.
- Extracción del año a partir de las fechas de aprobación para facilitar el análisis manual.

- Normalización del nombre de la universidad, dado que el dataset original contenía información adicional respecto a las sedes o facultades, lo cual dificultaba el filtrado de los datos por institución. Se creó una columna adicional donde se homogenizaron los nombres.
- Ocultar los datos personales de los beneficiarios, ya que aunque es información pública no es relevante para el estudio.

Una vez que se completaron estas acciones de limpieza, se decidió excluir el año 2024 del estudio, ya que se encontraba en curso y no disponía de información completa. Del mismo modo, se excluyó el año 2014, dado que presentaba montos significativamente menores en comparación con los años siguientes, lo que sugería que los datos podrían estar incompletos. Finalmente, se utilizó la información de los años 2015 a 2023 para el análisis.

## Análisis de monto y cantidad de beneficiarios por año

Se decidió graficar con todo el conjunto de datos, el comportamiento de la asignación de montos a lo largo de los años del dataset, para entender tendencias en la asignación de dinero. En la figura #2 se puede observar el comportamiento de las asignaciones de montos del 2015 al 2023.

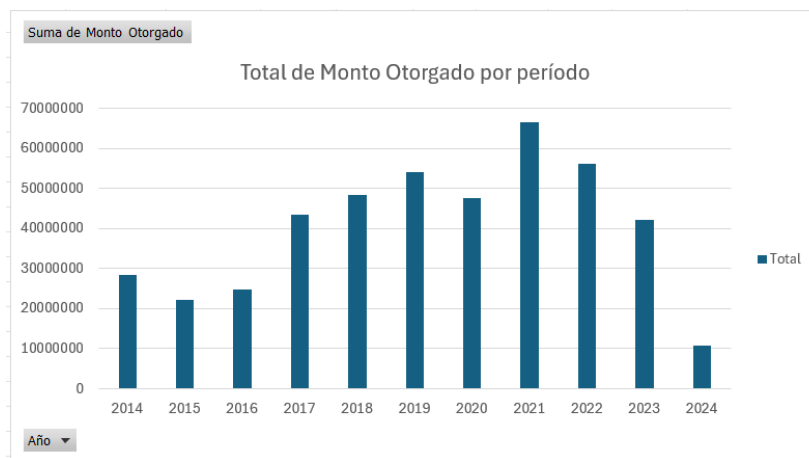


Figura #2: Gráfico total de monto otorgado por período. Elaboración propia.

En este gráfico se puede observar que a medida que van pasando los años, la tendencia es que aumente el monto otorgado por año.

Es importante recalcar que el dataset maneja 2 términos: monto otorgado y monto desembolsado. El monto otorgado es lo que se aprueba al solicitante mientras que el desembolsado es el entregado efectivamente durante el año en curso.

La gráfica presentada anteriormente puede contrastarse con la gráfica mostrada en la figura#3 donde se muestra total de beneficiarios por periodo.

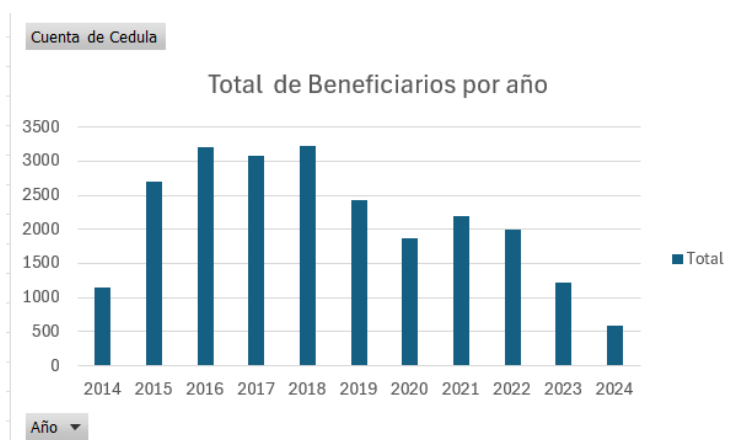


Figura #3: Total de Beneficiarios por año. Elaboración Propia.

Para explicar mejor la relación entre total de beneficiarios y total de monto aprobado, se decidió generar los siguientes indicadores:

- **Monto promedio por beneficiario:** permite entender cómo ha evolucionado la cantidad de dinero asignada a cada beneficiario. A lo largo de los años, el número de beneficiarios puede variar significativamente, pero es crucial evaluar si el monto otorgado por persona ha aumentado o disminuido.
- **Tasa de crecimiento del monto:** Este indicador mide el cambio porcentual en los montos otorgados de un año a otro. Puede ofrecer una clara perspectiva acerca de cómo ha fluctuado el presupuesto destinado a becas y préstamos a lo largo del tiempo.
- **Tasa de crecimiento de beneficiarios:** Este indicador muestra cómo ha cambiado el número de beneficiarios con el paso del tiempo, comparando el número de estudiantes que recibieron apoyo financiero en un año con el año anterior.

En la figura N°4 se observan los resultados obtenidos.

Etiquetas de fil	Suma de Monto Otorgad	Cuenta de Cedul	Monto Promedio Por Beneficiari	Tasa de Crecimiento del Monto:	Tasa de Crecimiento de beneficiario
2015	22083179.08	2699	8181.98558	12.65206979	18.59948129
2016	24877158.31	3201	7771.683321	74.42376183	-3.905029678
2017	43391675.36	3076	14106.52645	11.42724829	4.648894668
2018	48350149.84	3219	15020.23916	11.75613908	-24.38645542
2019	54034260.7	2434	22199.77843	-11.71570017	-23.29498767
2020	47703768.73	1867	25551.02771	39.40381257	17.72897697
2021	66500872.35	2198	30255.17395	-15.68644383	-8.780709736
2022	56069250.36	2005	27964.7134	-24.83708748	-39.50124688
2023	42143281.6	1213	34743.01863	-100	-100

FiguraN°4: Indicadores para el análisis del total de beneficiarios vs el total de monto otorgado. Elaboración propia

Las primeras columnas de la tabla corresponden a los años, total de monto otorgado por año y cantidad de beneficiarios. El primer indicador obtenido fue el monto promedio por beneficiario donde se puede observar una tendencia creciente en el

monto promedio, especialmente en los últimos años. Por ejemplo, en 2015 el promedio era de 8,181.98 dólares por beneficiario, mientras que, en 2023, el promedio subió significativamente a 34,743.02 dólares.

También se calculó la tasa del crecimiento del monto, donde se pueden observar fluctuaciones importantes. Por ejemplo, en 2016 hubo un aumento del 74.42% en el monto otorgado en comparación con 2015, mientras que en 2023, hubo una disminución del 100% en la tasa de crecimiento del monto, lo que podría ser un error o simplemente reflejar la caída en la cantidad de fondos disponibles o distribuidos ese año.

Otro indicador calculado es la tasa de crecimiento de beneficiarios, donde se indica el cambio porcentual en el número de beneficiarios en comparación con el año anterior.

A partir de 2019, la tasa de crecimiento de beneficiarios ha sido negativa, con una fuerte caída en 2023, con un -100% de tasa de crecimiento. Esto refleja una reducción drástica en la cantidad de beneficiarios durante este periodo.

## Predicción de monto Asignado a beneficiarios de la Universidad Tecnológica de Panamá

Basados en los análisis de datos realizados una vez el dataset estaba limpio, se pudo observar cierta estabilidad en los datos correspondientes a los montos otorgados a los beneficiarios de la Universidad Tecnológica de Panamá (UTP), tal como se muestra en la Figura #5.

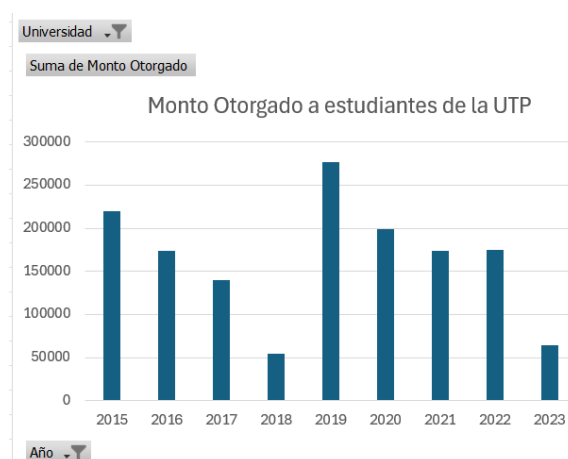


Figura #5: Gráfico de Montos otorgados a beneficiarios de la UTP. Elaboración propia.

Por ello se decidió evaluar el modelo con 4 periodos para ver si se obtenían resultados más estables, tal como se muestra en la figura N°7.



Periodo	Dt	Lt	Ft	Et	At	MSE	MAD	%Error	MAPE	Tst
1	219375.38									
2	173385.6									
3	139459.28									
4	54307.1	146631.84								
5	276454.52	160901.625	146631.84	-129822.68	129822.68	16853928242	129822.68	46.95986884	46.95986884	-1
6	199576	167449.225	160901.625	-38674.375	38674.375	9174817762	84248.5275	19.37826943	33.16906914	-2
7	174310.65	176162.0675	167449.225	-6861.425	6861.425	6132238226	58452.82667	3.936320013	23.42481943	-3
8	174951.08	206323.0625	176162.0675	1210.9875	1210.9875	4599545292	44142.36688	0.692186353	17.74166116	-3.94513264
9	64375.3	153303.2575	206323.0625	141947.7625	141947.7625	7709469689	63703.446	220.5003511	58.29339914	-0.50546292
									max	-0.50546292
									min	-3.94513264

Figura #7: Promedio Móvil con periodo de 4. Elaboración Propia.

Al usar periodos de 4, se visualizan mejores resultados en cuanto a los porcentajes de Error y MAPE. Se puede observar que en el período 5, el error es muy grande 129822.68 lo cual coincide con el período en la gráfica con el pico más alto.

Por lo tanto, se puede ver que los valores de Lt y el Ft son más estables.

### Suavización exponencial

El segundo modelo escogido para trabajar fue el modelo de Suavización Exponencial. Este modelo es utilizado para predecir el comportamiento futuro de una variable basándose en sus valores históricos. Este método asigna más peso a los datos más recientes, lo que permite una mejor adaptación a los cambios recientes en la serie temporal (Flandez, y otros, 2023).

Periodo	Dt	Lt	Ft	Et	At	MSE	MAD	%Error	MAPE	Tst
		164021.6567								
1	219375.38	172324.7152	164021.657	-55353.7233	55353.7233	3064034687	55353.7233	25.2324228	25.2324228	-1
2	173385.6	172483.8479	172324.715	-1060.88483	1060.88483	1532580082	28207.3041	0.61186444	12.9221436	-2
3	139459.28	167530.1627	172483.848	33024.5679	33024.5679	1385260749	29813.0587	23.6804377	16.5082416	-0.78455688
4	54307.1	150546.7033	167530.163	113223.063	113223.063	4243811044	50665.5597	208.486667	64.502848	1.77305892
5	276454.52	169432.8758	150546.703	-125907.817	125907.817	6565604497	65714.0111	45.5437722	60.7110328	-0.54896655
6	199576	173954.3444	169432.876	-30143.1242	30143.1242	5622771736	59785.5299	15.1035817	53.1097909	-1.10759106
7	174310.65	174007.7903	173954.344	-356.305565	356.305565	4819536767	51295.6407	0.20440837	45.5518792	-1.29785344
8	174951.08	174149.2837	174007.79	-943.28973	943.28973	4217205896	45001.5969	0.53917342	39.9252909	-1.50033595
9	64375.3	157683.1862	174149.284	109773.984	109773.984	5087552741	52198.5287	170.521899	54.4360251	0.80953374
					52198.5287					
									Max	1.77305892
									Min	-2
Alfa	0.15									

Figura N°8: Modelo de Suavización Exponencial. Elaboración propia.

Luego del cálculo de los valores, se estimó el uso de alfa en 0.15. Aquí se obtuvo valores interesantes como el MAD de 52198.53 y un MAPE de 54.43%. El rango de Tst se encuentra entre -2 y 1.7 lo cual indica menor tendencia a sobrestimar o subestimar valores en el modelo.

## Modelo de Holt

El modelo de Holt, también conocido como suavización exponencial doble, es una extensión del método de suavización exponencial simple, ya que puede capturar y predecir tendencias lineales en los datos. En la figura N°9 se muestra la aplicación del modelo de Holt (Rincón Rojas, 2021).

[illegible]

Figura N°9: Modelo de Holt. Elaboración Propia

El modelo de Holt tiene como componentes importantes el Alfa y el Beta los cuales influyen directamente en el cálculo de los valores del modelo. Se hizo varias pruebas con los valores a fin de encontrar los valores que mejor ajusten el modelo. Al final se utilizó un valor de alfa = 0.1 y beta = 0.15 lo cual trajo un nivel de Lt de ajuste más suave sin ser demasiado reactivo a cambios bruscos.

También se notan mejoras en los errores de los periodos, lo cuales indican el ajuste del modelo a los datos. Al igual que en el modelo de suavización exponencial se observan errores altos en los periodos 5 y 9 los cuales coinciden con los picos de los datos en los mismos periodos.

En cuanto al MSE este se presenta alto en periodos con grandes fluctuaciones como el periodo 5.

## Modelo ARIMA

El modelo ARIMA (Autoregressive Integrated Moving Average) es una técnica de análisis de series temporales utilizada para predecir valores futuros basándose en valores pasados. Es una combinación de tres componentes principales:

- Autoregresivo (AR): Utiliza la relación entre una observación y un número de observaciones anteriores.
- Integrado (I): Involucra la diferenciación de las observaciones para hacer que la serie temporal sea estacionaria.
- Media Móvil (MA): Modela el error de predicción como una combinación lineal de errores pasados.

En la figura N° 10 se observa los resultados del modelo ARIMA utilizando Python y Jupyter.

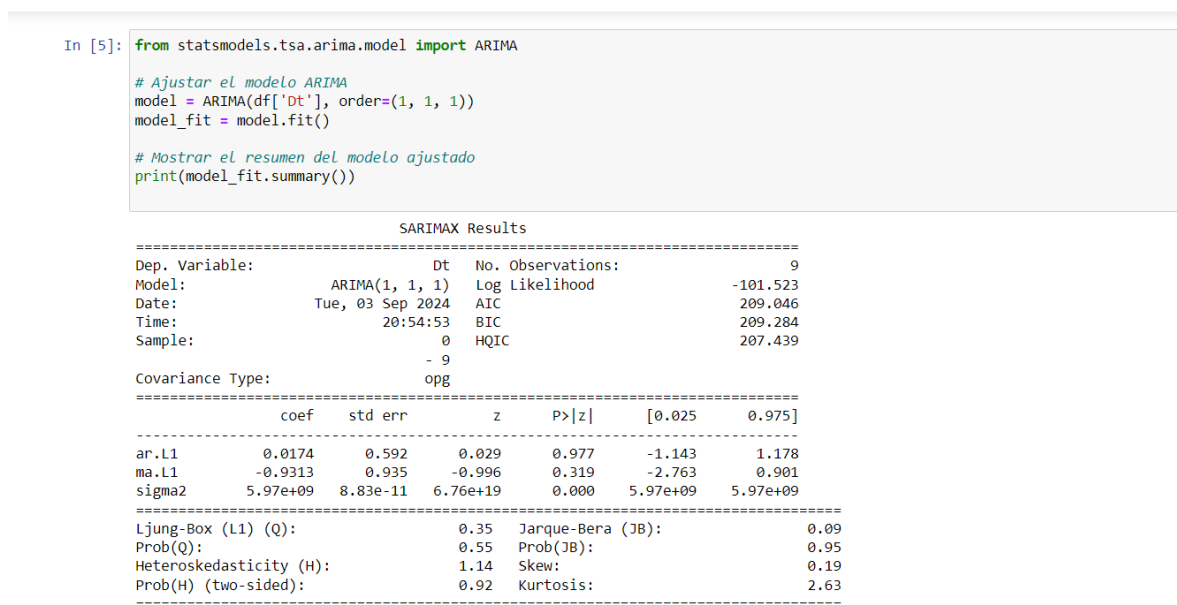


Figura N°10: Resultados del Modelo Arima.

El coeficiente AR(1) es muy pequeño y no significativo ( $P>|z|=0.977$ ). Esto indica que la dependencia autorregresiva (el uso de valores pasados para predecir el futuro) tiene poca influencia en este modelo.

El coeficiente MA(1) tiene un valor negativo (-0.9313), lo que significa que el error pasado tiene una influencia negativa sobre los valores futuros. Sin embargo, este coeficiente tampoco es significativo ( $P>|z|=0.319$ ).

Se generó una predicción para los siguientes 3 períodos:

- Periodo 10: 145817.85
- Periodo 11: 147237.65
- Periodo 12: 147262.40

Estas predicciones muestran una tendencia de estabilización o una leve variación hacia el alza en los siguientes periodos.

## Predicción de Cantidad de Beneficiarios de la Universidad Tecnológica de Panamá

Aprovechando la distribución de los datos y la duda obtenida del análisis de cantidad de beneficiarios vs monto asignado, se decidió también hacer la predicción respecto a la cantidad de beneficiarios en la UTP, usando los mismos métodos que en la predicción anterior, sin tomar en cuenta los años 2014 y 2024. En la figura N°11 pueden observarse los beneficiarios de la UTP en los periodos 2014-2024.

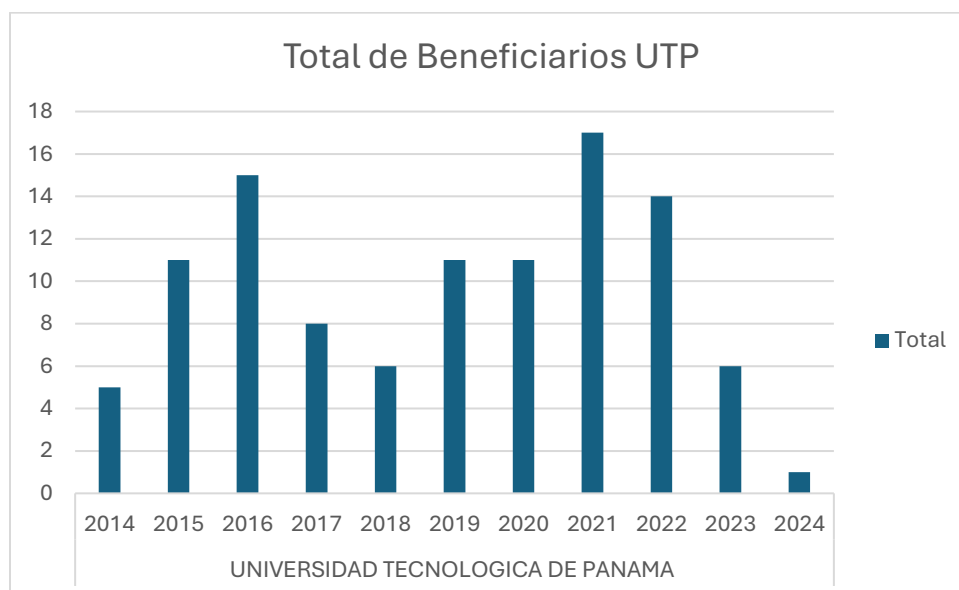


Figura N°11: Total de Beneficiarios de la UTP, entre 2014-2024. Elaboración propia.

## Promedio Móvil

Al igual que en la predicción anterior, se comenzó con el modelo de promedio móvil, utilizando un periodo de 3 años para realizar el análisis tal como se observa en la figura N°12

Periodo	Dt	Lt	Ft	Et	At	MSE	MAD	%Error	MAPE	TSt
1		11								
2		15								
3		8								
4		6	11.3333333							
5		11	9.66666667	11.3333333	0.33333333	0.33333333	0.111111111	0.33333333	3.030303	1
6		11	8.33333333	9.66666667	-1.33333333	1.33333333	0.944444444	0.83333333	12.12121	-1.2
7		17	9.33333333	8.33333333	-8.66666667	8.66666667	25.66666667	3.44444444	50.98039	-2.80645
8		14	13	9.33333333	-4.66666667	4.66666667	24.69444444	3.75	33.33333	-3.82222
9		6	14	13	7	7	29.55555556	4.4	116.6667	-1.66667
				14						
									max	1
									min	-3.82222

Figura N°12: Promedio móvil. Periodo de 3 años. Elaboración Propia

Se puede observar que el valor de MAPE de 43.23% en el último período indica que hay una variabilidad considerable en el ajuste del modelo a los datos, lo que sugiere que el promedio móvil de 3 períodos no es el ajuste más preciso. Otro indicador clave son los valores de Tst, especialmente el valor negativo de -3.82, el cual indica que el pronóstico puede ser más alto de lo que debería hacer. El porcentaje de error es elevadamente alto en el último periodo.

Así que, en vista de esto, se decidió experimentar con un período de 4 años, intentando encontrar mejor precisión del modelo, tal como se muestra en la figura N°13.

Periodo	Dt	Lt	Ft	Et	At	MSE	MAD	%Error	MAPE	TSt
1		11								
2		15								
3		8								
4		6	10							
5		11	10	10	-1	1	1	1	9.09090909	-1
6		11	9	10	-1	1	1	1	9.09090909	-2
7		17	11.25	9	-8	8	22	3.33333333	47.0588235	-3
8		14	13.25	11.25	-2.75	2.75	18.390625	3.1875	19.6428571	-4
9		6	12	13.25	7.25	7.25	25.225	4	120.833333	-1.375
									max	-1
									min	-4

Figura N°13: Promedio móvil. Periodo de 4 años. Elaboración Propia

En este caso se observan algunas variaciones por ejemplo que el MAPE se reduce a 41.143. En cuanto al porcentaje de error, este aumentó en el último periodo. Y en cuanto a los valores del Tst se observa que todos están en el rango negativo, con el

El modelo de Holt también fue implementado en este caso, donde en primer lugar se hace la regresión lineal de los valores de periodo y dt. En la figura N°15 se aprecia la aplicación del modelo de Holt.



SARIMAX Results						
=====						
Dep. Variable:		Dt	No. Observations:			9
Model:	ARIMA(1, 1, 1)		Log Likelihood			-23.149
Date:	Fri, 06 Sep 2024		AIC			52.299
Time:	20:35:00		BIC			52.537
Sample:	0		HQIC			50.691
	- 9					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
ar.L1	0.2091	0.937	0.223	0.823	-1.628	2.046
ma.L1	-0.9999	1505.092	-0.001	0.999	-2950.927	2948.927
sigma2	15.1908	2.29e+04	0.001	0.999	-4.48e+04	4.48e+04
=====						
Ljung-Box (L1) (Q):			0.00	Jarque-Bera (JB):		0.52
Prob(Q):			0.97	Prob(JB):		0.77
Heteroskedasticity (H):			1.57	Skew:		-0.05
Prob(H) (two-sided):			0.72	Kurtosis:		1.75
=====						
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						
Pronósticos para los próximos periodos:						
9	9.844670					
10	10.648655					
11	10.816782					
Name: predicted_mean, dtype: float64						

Figura #16: Modelo ARIMA. Elaboración Propia.

Dentro de los pronósticos obtenidos con ARIMA se observa que los periodos 9, 10 y 11, están en el mismo rango con los valores anteriores, lo cual sugiere valores más estables luego de la fluctuación del periodo 8.

En cuanto a las métricas, ARIMA arroja un MSE de 15.19, con una precisión aceptable, algo mejor que el resultado obtenido por Holt. En el MAD ARIMA brinda un valor de 3.00 donde también muestra mayor precisión que Holt.

## Resultados

Por sugerencia del profesor, se hizo un análisis para profundizar las razones en porque los datos sugieren que a través de los años aumentó la cantidad de monto otorgado mientras que fue disminuyendo la cantidad de beneficiarios

A lo largo de los años, aunque el número de beneficiarios ha disminuido, el monto promedio otorgado por beneficiario ha aumentado considerablemente, lo que indica que la institución está destinando mayores recursos por persona. En el 2023 se muestra una caída importante en las tasas de crecimiento tanto de montos como de número de beneficiarios, esto es interesante ya que este año fue cuando comenzó la polémica pública sobre las formas en que se concedían los beneficios del programa Auxilios Económicos, que ocupan buena parte de los datos del data set utilizado en este proyecto.



Para la predicción de la cantidad de monto que será otorgada a los estudiantes de la Universidad Tecnológica de Panamá, en la figura N° 17 se presenta la comparativa de resultados entre los diversos métodos probados

ESTIMACIONES DE ERROR DE LOS PRONÓSTICOS					
Método	MAD	MAPE	Rango TS inf	Rango TS sup	Desv. Est.
Promedio Movil (P=3)	69608.668	236.7446676	-3.92524461	-0.56856291	87,011
Promedio Movil (P=4)	63703.446	220.5003511	-3.94513264	-0.50546292	79,629
Suavización Expo	52198.52874	54.43602511	-2	1.77305892	65,248
Holt	55526.63547	53.70224135	-1.56560728	2.83457957	69,408
Arima	61877.00134	54.89312902	-0.148872	3.938415	77,346

Figura N°17: Estimaciones de error de los pronósticos. Elaboración propia

A continuación, se presenta el análisis de los resultados para cada método:

### Error Absoluto Medio (MAD)

El método de Suavización Exponencial resultó ser el método más preciso con un MAD de 52198.58, lo que indica que, en promedio, los errores de pronóstico son más bajos en comparación con otros métodos. Por otro lado, el Promedio Móvil con P=3 mostró el MAD más alto, lo que sugiere que este método no es el más adecuado para los datos analizados.

### Error Porcentual Medio Absoluto (MAPE)

El método de Holt tuvo el promedio de MAPE más bajo 53.70 % seguido de Suavización Exponencial 54.43% y ARIMA 54.89%.

### Rango de la Señal de Seguimiento (TS)

El modelo ARIMA muestra el Rango TS más estable, con valores cercanos a 0, con un valor de -0.148872 a 3.938415, lo que sugiere que sus predicciones son las más consistentes y con menos sesgo.

### Desviación Estándar

La Suavización Exponencial mostró la Desviación Estándar más baja (65,248), indicando una mayor consistencia en sus errores.

Los modelos de ARIMA y Holt presentaron desviaciones estándar similares, de 77,346 y 69,408 respectivamente, lo que indica una consistencia moderada en sus errores.

Por lo tanto, se puede concluir que el método de Suavización Exponencial ofrece el mejor desempeño general, lo cual es compatible con su aplicabilidad a datos sin tendencia ni estacionalidad. En segundo lugar, puede recomendarse el método ARIMA.

Así que, el monto que otorgará el IFARHU a los estudiantes de la UTP durante el periodo 10 (año 2024) será de 157683.18 dólares

Para el análisis de cantidad de beneficiarios de la Universidad Tecnológica de Panamá, en la figura #18 se presenta la comparativa de resultados entre los modelos utilizados.

ESTIMACIONES DE ERROR DE LOS PRONÓSTICOS					
Método	MAD	MAPE	Rango TS inf	Rango TS sup	Desv. Est.
Promedio Movil (P=3)	4.4	116.6666667	-3.82222222	1	6
Promedio Movil (P=4)	4	120.8333333	-4	-1	5
Suavización Expo	3.278100296	36.66959673	-2	1.50109327	4
Holt	3.828805076	42.95337359	-2.20076912	1.68325792	5
ARIMA	2.014595985	22.6325095	-1.916993	1.687116	3

Figura N°18: Estimaciones de error de los pronósticos. Elaboración propia

### **MAD (Desviación Media Absoluta)**

El modelo de Suavización Exponencial tiene un MAD de 3.28, mostrando una reducción significativa del error comparado con el Promedio Móvil, lo que indica un ajuste más preciso.

El modelo ARIMA, con un MAD de 2.01, es el modelo que muestra la menor desviación promedio. Esto indica que sus predicciones son las más cercanas a los valores reales, destacándose como el modelo más preciso en esta métrica.

En resumen, El modelo ARIMA es el más preciso en términos de MAD, seguido de cerca por Suavización Exponencial. Ambos modelos son superiores al Promedio Móvil y Holt, que muestran errores más elevados.

### **MAPE (Error Porcentual Absoluto Medio)**

El método de Suavización Exponencial presenta un MAPE de 36.67%, lo que indica que aunque todavía hay proporción considerable de error,

ARIMA tiene el MAPE más bajo, con 22.63%, lo que significa que sus errores porcentuales son menores y más consistentes en relación con los valores reales. Este valor demuestra que ARIMA es el más efectivo en la predicción precisa de los datos.

### **Intervalo de Tracking Signal (TS)**

ARIMA tiene el rango de TS más estrecho, con valores entre -1.92 y 1.68, lo que indica que sus predicciones son las más consistentes y menos sesgadas. Esto sugiere que ARIMA es el modelo más equilibrado, con menor tendencia a sobrestimar o subestimar los valores.

### **Desviación Estándar**

ARIMA presenta la desviación estándar más baja con un valor de 3, lo que indica que sus errores son los más consistentes. Esto significa que ARIMA no solo tiene errores más bajos, sino que también es el más confiable en cuanto a la estabilidad de las predicciones.

Por lo tanto, se puede afirmar que el modelo ARIMA demuestra ser el más preciso y estable, con los valores más bajos de MAD, MAPE, y Desviación Estándar, además de un rango de TS más estrecho. Esto lo convierte en la opción más confiable para predecir los montos futuros en este conjunto de datos. Si hubiera que escoger una segunda opción, sería el método de suavización exponencial, que en todos los criterios ocupó el segundo lugar respecto a ARIMA.

Por lo tanto, la cantidad de estudiantes beneficiados por los programas del IFARHU en la UTP durante el periodo 10 (año 2024) será de 9.85 es decir 10 beneficiario.

## Hallazgos

Aquí presentamos algunos hallazgos o situaciones que nos llamaron la atención, mientras se analizaban los datos y se aplicaban los modelos predictivos:

- Hay indicios que la información del dataset está incompleta. Esto debido a que analizando los datos de los beneficiarios de la UTP se ven datos muy pequeños en las cantidades, lo cual resulta llamativo, tomando en cuenta que la UTP es una universidad con presencia a nivel nacional con más de 20mil estudiantes (Datos Abiertos de Panamá, s.f.).
- Cuando se publicó el dataset utilizado para este estudio, los representantes del IFARHU dijeron que aquí se encontraría la información referente a los auxilios económicos. Sin embargo, se ha podido ver en la data que hay más programas incluidos. Además, en los medios de comunicación se han presentado denuncias de personas que solicitaron otro tipo de apoyos (becas, préstamos) y luego salieron en el listado de auxilios económicos.
- El data set utilizado no contempla información acerca de las becas de concurso de todos los niveles, así que la cantidad de beneficiarios y cantidad de dinero debe ser mayor a la mostrada en el dataset. Si esta información hubiera estado completa tal vez se hubiera nivelado la gran diferencia entre montos que había entre las universidades estatales.

## Conclusiones

Luego de haber realizado el análisis de los datos, la limpieza, la aplicación de los modelos y la revisión de los resultados, se pueden concluir varias ideas importantes:

- Es de vital importancia contar con datos correctos, ordenados, completos que ofrezcan una muestra real de la información que se desea predecir. Si la fuente de datos no está correcta, de salida todo el proceso obtendrá resultados inexactos.
- Se deben conocer las bases estadísticas para la manipulación de los datos y la lectura de los resultados de los modelos.
- Es importante conocer que tipo de resultados brindan los modelos y como se deben interpretar.
- Tener conciencia de que a las finales las predicciones son solo eso, una valoración de lo que podrá ocurrir tomando en cuenta los datos históricos. Es posible que las predicciones fallen por múltiples situaciones, pero lo que hay que tener seguridad es que el proceso de los modelos predictivos se hizo de forma correcta y se aplicaron las técnicas necesarias para evitar sesgos, malas interpretaciones o manejo inadecuado de las métricas presentadas por los modelos.

## Bibliografía

*Datos Abiertos de Panamá.* (s.f.). Obtenido de <https://www.datosabiertos.gob.pa/dataset/utp-datos-de-matricula-por-sexo-y-sede-ano-academico-2022-2023>

Flandez, J., Vargas, R., Jacques, V., Cornejo, M., Gajardo-Burgos, R., Giakoni-Ramirez, F., . . . Duclos-Bastías, D. (2023). Rendimiento en los Juegos Panamericanos de Santiago 2023: Predicción mediante método de suavización exponencial simple. *Revista Ciencias de la Actividad Física UCM*, 24.

*IFARHU.* (s.f.). Obtenido de IFARHU: <https://www.ifarhu.gob.pa/institucional/mision-y-vision/>

Prades Escobar, E., & Marín Sánchez, D. (2020). Modelos estadísticos para las predicciones de la COVID-19 en Cuba. *Revista Cubana de Higiene y Epidemiología*, 57.

Rincón Rojas, D. (2021). Modelo predictivo para determinar la tendencia de los casos de Covid-19 mediante la técnica de suavizado exponencial doble. *Universidad Distrital Francisco José de Caldas - Biblioteca UDFJC*.

