

## **INFORME DEL PROYECTO:**

TECNICATURA SUPERIOR EN CIENCIAS DE DATOS E INTELIGENCIA ARTIFICIAL

**ESPACIO CURRICULAR:** ANALISTA DE DATOS I

**PROF.** UGARTE MARCOS // PERATTA NAHUEL

### **AMERICAN THINGS – ANALISIS Y PROCESAMIENTO DE DATOS, PARA SELECCIÓN DE ESTRATEGIAS DE VENTAS.**

**Introducción y Desarrollo:**

**AÑO 2025**

**Integrantes:**

Guaraz Emanuel - <https://github.com/JEmanuelG>

Direni Carlos - <https://github.com/Cdireni1>

Ryser Lucas - <https://github.com/lucasryser6>

Lobo Bruno - <https://github.com/lobosouza>

Allende Olmedo Nicolás - <https://github.com/AllendeNicolas>

Lugar de presentación: ISPC (Instituto Superior Politécnico de Córdoba)

# INTRODUCCIÓN:

## 1. NOMBRE DEL PROYECTO:

American Things – análisis y procesamiento de datos, para selección de estrategias de ventas.

Git hub del Proyecto: <https://github.com/AllendeNicolas/Analista-de-datos-I/tree/main>

## 2. TIPO DE PROYECTO:

El proyecto tiene un enfoque ***De Investigación y tecnológico***, donde se propone el análisis y procesamiento de datos, relacionados a las ventas más actuales de la empresa, obtenidos en un período de tiempo, con el fin de desarrollar y aplicar, tecnologías y métodos destinados a poner en conocimiento al establecimiento, de las categorías de productos, que han tenido un alza o una baja en ventas, y a partir de ello, abordar las estrategias necesarias para mejorar los diferentes sectores y generar un mejor rendimiento.

***Estratégico y de Marketing***, ya que apuntamos a que la propuesta del proyecto sea, el puntapié inicial, de una serie de toma de decisiones de marketing y publicidad, que apunten a generar, una mejora en las ventas de cada una de sus categorías, y reforzar aquellas, cuyas ventas son prometedoras.

***De Investigación y Procesamiento***, mediante el análisis a aplicación de software, utilizados en la materia Procesamiento de datos, y las técnicas en exploración y estadísticas, conocer y explotar de manera eficiente, todos los datos brindados, realizando una limpieza y puesta a punto del “Dataset”, y utilizar cada uno de esos aportes, para seleccionar las mejores estrategias de venta, y determinar las decisiones más acertadas, relacionadas con el marketing de cada categoría de dicho establecimiento.

## 3. ESPACIO CURRICULAR:

Analista de Datos I – Profesor/es: Marcos Ugarte y Nahuel Pratta.

#### **4. EJES TEMÁTICOS/RED DE CONCEPTOS:**

En el contexto del módulo de Análisis de Datos, el presente proyecto responde a la necesidad de la empresa American Things de llevar a cabo análisis estratégicos eficaces. Este trabajo se centra en las etapas del procesamiento de datos, con el propósito de asegurar la calidad de los datos a analizar, así como en las técnicas de exploración de datos y en los métodos estadísticos empleados para obtener resultados precisos y confiables. El proyecto integra los contenidos de las materias Procesamiento de Datos y Estadística y Exploración de Datos I bajo el siguiente enfoque metodológico:

##### **1. Identificación y Preparación del Conjunto de Datos:**

- Aplicación de técnicas de limpieza
- Normalización de los datos
- Transformación en dataset utilizando herramientas como Python (pandas) y/o SQL, para garantizar la calidad y coherencia de los datos.

##### **2. Definición de Variables y Exploración:**

- Clasificación de variables
- Organización de los datos
- Presentación de los datos en tablas y gráficos
- Cálculo de medidas numéricas

##### **3. Detección de Tendencias y Segmentación Estratégica**

- Interpretación de la información obtenida.

#### **5. PROBLEMÁTICAS/NECESIDADES:**

American Things, un poli rubro, dedicado a la ventas de productos en general, se ha visto en la necesidad de organizar la información alojada en sus base de datos, respecto a las ventas realizadas por esta, en un determinado periodo de tiempo. Su Data Set, presenta anomalías respecto de los datos que han sido cargados en ella, ya sea por un error externo, falta de previsión al ingresar datos a la base por sujetos no pertenecientes a la empresa, o internos, por no tomar los cuidados y las diligencias necesarias al momento de trabajar con la información y los datos proporcionados en la carga del Dataset, por personas encargadas pertenecientes al mismo establecimiento. La información incompleta o mal dispuesta de la base de datos, no permite a American Things, realizar análisis estratégicos eficientes, para tomar decisiones efectivas, sobre las distintas categorías de productos que ofrecen, y no les permite conocer con precisión, cuales son los sectores que deben atender y mejorar, para lograr los resultados de ventas deseados, de manera general y específica en la empresa.

## 6. FUNDAMENTACIÓN:

La elección del problema, al cual está orientado el proyecto, se centra en la necesidad persistente de una constante actualización, mantenimiento y análisis, de las diversas bases de datos que una empresa pueda tener.

En un entorno empresarial altamente competitivo, el manejo eficiente de los datos es un pilar fundamental para la toma de decisiones estratégicas.

### **Impacto de una base de datos inconsistente:**

Las inconsistencias en la base de datos generan problemas que repercuten en varios aspectos clave:

- **Dificultad en la segmentación de clientes:** Sin datos precisos, las empresas no puede identificar patrones de compra ni definir estrategias de marketing efectivas.
- **Imposibilidad de análisis de ventas preciso:** La empresa no puede determinar con exactitud qué productos tienen mayor demanda en determinados períodos.
- **Pérdida de oportunidades de optimización:** La toma de decisiones basada en datos defectuosos puede llevar a inversiones erróneas o falta de atención a segmentos estratégicos.
- **Riesgos operativos:** Información incorrecta puede generar problemas logísticos, afectando la gestión de inventario y el abastecimiento oportuno de productos.

En Conclusión, el manejo eficiente de la base de datos es un factor determinante para la competitividad de American Things en el mercado. La depuración y control de la información permiten generar estrategias basadas en datos confiables, optimizar la gestión de ventas y potenciar la toma de decisiones estratégicas. Implementando medidas correctivas y preventivas, la empresa puede garantizar un sistema de información robusto, con datos precisos y útiles para su crecimiento sostenido.

## 7. VISIÓN DEL PROYECTO:

### **OBJETIVO GENERAL:**

- Ordenamiento, Limpieza, Exploración y Análisis de datos, para mitigar problemas y asegurar la fiabilidad del Dataset, para implementación de medidas estratégicas.

### **OBJETIVOS ESPECÍFICOS:**

1) Analizar costos fijos y gastos variables. (Aquellos que puedan surgir por la ejecución de la iniciativa de innovación y gestión de organización de la base de datos).

- 2) Seleccionar los medios necesarios, software y hardware, de aplicación. (Aquellos medios que nos permitan realizar la limpieza y organización del Dataset, sin dañar y cuidado los datos existentes).
- 3) Desarrollar un entorno seguro de ejecución. (Que se tomen todas las precauciones y medidas de seguridad, para evitar abusos, estafas, y manipulación ilegal de los datos).
- 4) Validación y depuración periódica del Data Set. (Mediante técnicas de limpieza de datos para corregir registros incorrectos o duplicados).
- 5) Protocolos estrictos de ingreso de información. (Definir procesos estandarizados que eviten errores humanos en la carga de datos).
- 6) Capacitación del personal. (Asegurar que quienes manipulan la base de datos comprendan la importancia de la calidad de la información y sigan buenas prácticas de gestión).
- 7) Control de acceso y permisos. (Limitar el ingreso de información solo a personal autorizado para prevenir manipulaciones externas no controladas).
- 8) Implementación de herramientas de gestión automatizadas. (Utilizar software avanzado para verificar la integridad y coherencia de los datos, evitando errores manuales).

#### **METAS:**

- 1) Implementar el Dataset, actualizado, limpio y ordenado, para una explotación de datos más eficiente, y así aplicar una mejor calidad de información, a las nuevas decisiones estratégicas de mercado. (Utilización del Dataset. Aplicado los nuevos resultados del análisis de la base de datos, más fiables y óptimos).

#### **8. SELECCIÓN DE ACCIONES:**

<i><b>Objetivo general</b></i>	<i><b>Acciones</b></i>	<i><b>Habilidades o capacidades a lograr</b></i>
Ordenamiento, Limpieza, Exploración y Análisis de datos.	Mitigar problemas y asegurar la fiabilidad del Dataset, para implementar las medidas estratégicas pertinentes.	Gestionar y aplicar, medidas y técnicas, tendientes a la limpieza y conservación de los datos, y asegurar la fiabilidad del Dataset.

<i><b>Objetivos específicos</b></i>	<i><b>Acciones</b></i>	<i><b>Habilidades o capacidades a lograr</b></i>
Analizar costos fijos y gastos variables.	Desglosar e identificar, de manera específica, cada gasto y costo, que pueda suscitarse, en la modificación y actualización del Dataset, por el uso de Software y tecnologías destinadas a tal fin.	Determinar con exactitud, cuales son aquellos gastos y costos a considerar, para el desarrollo, y gestión de la Base de datos, y posteriormente, de su aplicación.
Seleccionar los medios necesarios, software y hardware de aplicación.	Conocer y aplicar, aquellos medios que nos permitan realizar la limpieza y organización del Dataset, sin dañar y cuidando los datos existentes.	Elegir los métodos, técnicas, software y demás tecnologías, que brinden una manipulación de los datos, de una manera, fluida y sencilla
Desarrollar un entorno seguro de ejecución.	Tomar todas las precauciones y medidas de seguridad, para evitar abusos, estafas, y manipulación ilegal de los datos	Determinar y obtener las autorizaciones necesarias, para la manipulación del Dataset, cuidando de proteger, todos aquellos datos que sean sensibles, o que requieran una atención especial de protección.
Validación y depuración periódica del Data Set.	Seleccionar técnicas de limpieza de datos para corregir registros incorrectos o duplicados.	Asegurar la limpieza del Dataset, rellenar datos faltantes, eliminar datos duplicados, recuperar datos que se hayan perdido y modificar aquellos mal ingresados, entre otras acciones.
Protocolos estrictos de ingreso de información.	Definir procesos estandarizados que eviten errores humanos en la carga de datos.	Crear un protocolo, que asegure la carga de datos de manera eficiente, evitando, en lo posible, la mayor cantidad de errores posibles.
Capacitación del personal.	Asegurar que quienes manipulan la base de datos, comprendan la importancia de la calidad de la información y sigan buenas prácticas de gestión.	Brindar la información necesaria, al personal autorizado, sobre los protocolos de ingresos de datos.

Control de acceso y permisos.	Limitar el ingreso de información solo a personal autorizado para prevenir manipulaciones externas no controladas.	Seleccionar al personal, encargado de realizar, la carga y manipulación de los datos, de cada Dataset de la empresa.
Implementación de herramientas de gestión automatizadas.	Utilizar software avanzado para verificar la integridad y coherencia de los datos, evitando errores manuales.	Aplicar software, que sirva de apoyo y guía, a aquellos que estén autorizados a gestionar la Base Datos, para una eficaz manipulación de esta.

## 9. CRONOGRAMA:

<b>CRONOGRAMA</b>	<b>MES 0</b> <b>Semana</b> <b>01-02</b>	<b>MES 0</b> <b>Semana</b> <b>03-04</b>	<b>MES 1</b> <b>Semana</b> <b>01-02</b>	<b>MES 1</b> <b>Semana</b> <b>03-04</b>
Ordenamiento, Limpieza, Exploración y Análisis de datos	Diagnóstico del Dataset actual  Selección y limpieza inicial de datos	Normalización y transformación de datos  Exploración y primeros análisis	Análisis de tendencias Generación de reportes	Revisión final de análisis  Preparación de informe de estrategias
Selección de herramientas	Relevamiento de herramientas necesarias  Evaluación de alternativas	Adquisición e instalación de software y hardware	Capacitación en el uso de herramientas seleccionadas	Implementación de control y gestión

Desarrollo de un entorno seguro de ejecución	Diagnóstico de seguridad actual. Recomendaciones iniciales.	Diseño de protocolos de acceso y permisos	Capacitación en protocolos de seguridad	Aplicación y auditoría de medidas de seguridad
Validación y depuración periódica del Dataset	Primera depuración de datos (detección de errores)	Segunda depuración y validación cruzada	Corrección de inconsistencias finales	Validación definitiva del Dataset
Protocolos estrictos de ingreso de información	Diseño preliminar de protocolos	Revisión y ajuste de protocolos con feedback interno	Capacitación del personal	Implementación formal de protocolos
Capacitación del personal	Planificación de capacitaciones	Ejecución de capacitaciones técnicas	Ejecución de capacitaciones sobre buenas prácticas	Seguimiento y evolución del personal capacitado
Control de acceso y permisos	Identificación de roles críticos	Definición de niveles de acceso	Asignación de permisos	Revisión de acceso y ajustes
Implementación de herramientas de gestión automatizadas	Selección de herramientas de gestión	Implementación inicial	Integración con base de datos	Optimización y ajustes finales

**Presupuesto con gastos fijos y variables.**

*Al mes 5 de 2025 - tipo de cambio \$1200 ARS por dólar oficial.*



Son los gastos que se van a mantener durante el proceso que dure el proyecto o que se contratarán por única vez.

### **COSTOS FIJOS**

Concepto	Detalle	Estimado USD	Estimado \$ (pesos argentinos)
Comunicación Interna	Plataforma de comunicación Suscripción anual a herramienta tipo Slack o Microsoft Teams (10 usuarios)	100	120.000
Software de análisis de datos	Python Python (con librerías Pandas, NumPy, Matplotlib) y Excel, Google Sheet.	15.00 Excel / Python Licencia Libre	18.000
Servicios en la nube para almacenamiento	Google o Dropbox (2 TB)	9.99/mes	12.000
Hardware	Notebook de gama media (i5, 16GB RAM, SSD)	1.200	1.440.000
Capacitación de personal	curso básico de limpieza de datos	50	60.000 x curso
Software de gestión y visualización de datos	Licencia Power BI Pro	120/año	144.000

### **GASTOS VARIABLES**

Concepto	Detalle	Estimado USD	Estimado \$
Hs. consultoría técnica	Consultor de datos freelance (20 hs)	200.00	240.000
Servicios de seguridad informática	Instalación	50.00	60.000
Revisión y auditoría	Validación de la limpieza final del	100.00	120.000

externa	dataset.		
Eventos de capacitación interna	Reuniones, charlas de protocolos de carga de datos.	30.00 (catering)	36.000
Publicidad	Diseño Gráfico, Flyers Digitales, Campañas de email	500	600.000

## 10. PRODUCTO FINAL:

Un software de análisis y procesamiento de datos relacionado con las ventas de una empresa, recibiendo como entrada un dataset con información bruta, la cual sería procesada mediante una limpieza y normalización de los datos utilizando fundamentalmente la librería Pandas del lenguaje Python, entre otras como NumPy, dando como resultado información sobre:

- Altas y bajas en las ventas de determinados rubros o categorías en distintos rangos de fechas.

- Promedios de valoraciones de productos.

El proyecto se lleva a cabo gracias a la participación y acompañamiento del espacio curricular “Analista de Datos”, donde adquirimos los conocimientos y herramientas para poder llevar a cabo dicho proyecto.

## 11. BIBLIOGRAFÍA:

Triola, M. F. (2009). *Estadística*. Pearson Educación

Mckinney, W. (2017). *Python Para Análisis De Datos*. Anaya Multimedia.