**INSTITUTO SUPERIOR POLITÉCNICO DE CÓRDOBA**

**TECNICATURA SUPERIOR EN CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL**

**Módulo de Práctica Profesionalizante**

**Entrega 1**

**Introducción y onboarding**

**26/08/2024**

**Integrantes:**

* **López, Erick**
* **Nüesch, Christian**

**Índice**

[**Desarrollo de la Idea de Negocio**](#_cetjbdnzym28) **3**

[Idea principal](#_qd486u8pu59v) 4

[Propuesta de valor](#_wuwvetrkpred) 5

[**Expectativas y Objetivos del Proyecto**](#_yd1k5wyp8s2j) **6**

[Objetivos a corto plazo:](#_ozn880s1hqsr) 6

[Objetivos a largo plazo:](#_rj125380ju94) 8

[Objetivos del Equipo:](#_k2plbkz6f20p) 9

[**Metas Individuales**](#_zgn327lo71ii) **9**

[Roles y Responsabilidades](#_fd8l40cmjgts) 9

[**Plan de Acción**](#_ti0cn76rosuk) **12**

[Fases del Proyecto](#_mc6s4rqia692) 12

[Hitos y Entregables](#_lgrj6u303pwc) 13

[Metodología de Trabajo](#_q2stnyp910t6) 14

# Desarrollo de la Idea de Negocio

DataVista Analytics es una empresa especializada en análisis de datos, y ha

identificado varias oportunidades de negocio basadas en el análisis del dataset df\_autos. Este dataset, recopilado a principios de 2023, incluye información detallada sobre ventas de automóviles, como marca, modelo, tipo de chasis, precio de venta y la moneda en la que se cotiza (pesos o dólares).

La riqueza y variedad de estos datos proporcionan una base sólida para la generación y desarrollo de nuevas ideas de negocio. Al analizar este dataset, podemos obtener insights valiosos y detectar oportunidades en el mercado automotriz, facilitando la identificación de tendencias y áreas de oportunidad que pueden ser explotadas para el crecimiento y éxito de nuestra empresa.

A continuación, se presentan cuatro ideas principales de negocio que podrían desarrollarse utilizando este dataset:

* **Informes de mercado personalizados**:  
  Idea: Ofrecer informes a medida para concesionarios, fabricantes o empresas de seguros, basados en el análisis del dataset **df\_autos**.  
  Ejemplo: Un informe podría analizar las preferencias de los clientes en cuanto a marcas, modelos o características de seguridad, ayudando a un concesionario a optimizar su inventario.
* **Predicción de precios**:  
  Idea: Desarrollar un modelo de Machine Learning que prediga el precio de un automóvil usado basándose en las características del dataset.  
  Ejemplo: Ofrecer este modelo como un servicio a sitios web de compraventa de automóviles o a empresas de tasación.
* **Detección de anomalías**:  
  Idea: Analizar el dataset para identificar posibles fraudes o inconsistencias en los datos de los vehículos.  
  Ejemplo: Detectar anuncios de autos con kilometraje sospechosamente bajo o precios inusualmente altos.
* **Segmentación de clientes**:  
  Idea: Agrupar a los clientes potenciales en función de sus preferencias automotrices utilizando técnicas de clustering.  
  Ejemplo: Dirigir campañas de marketing personalizadas a cada segmento, ofreciendo vehículos que se ajusten a sus necesidades.

## Idea principal

En base a los datos disponibles, la opción con mayores posibilidades de implementación es la de “Informes de mercado personalizados”.

**Problema que se aborda:** En el competitivo mercado automotriz, concesionarios, fabricantes y empresas de seguros enfrentan el desafío de tomar decisiones informadas en un entorno dinámico y en constante cambio. La falta de acceso a datos precisos y análisis detallados puede llevar a decisiones subóptimas, afectando la rentabilidad y el crecimiento.

**Solución propuesta:** DataVista Analytics propone la creación de informes de mercado personalizados, basados en el análisis exhaustivo del dataset df\_autos. Estos informes proporcionarán insights valiosos sobre tendencias del mercado, preferencias de los clientes, análisis de la competencia y oportunidades de mercado. La solución incluye:

* Identificación de Necesidades de los Clientes Potenciales:
  + Determinar qué tipo de información requieren los concesionarios, fabricantes y empresas de seguros.
  + Identificar las preguntas clave que pueden responderse utilizando los datos del dataset.
* Definición de los Tipos de Informes a Ofrecer:
  + Tendencias del mercado (marcas, modelos, precios).
  + Preferencias de los clientes (características, tipo de vehículo).
  + Análisis de la competencia.
  + Oportunidades de mercado.
* Desarrollo de Herramientas de Análisis:
  + Escribir funciones en Python para procesar y analizar los datos del dataset.
  + Utilizar librerías como Pandas y Matplotlib para generar visualizaciones en forma de gráficos y tablas.
* Creación de Plantillas de Informes:
  + Diseñar informes con una estructura clara y visualmente atractiva.
  + Incluir gráficos, tablas y resúmenes de los hallazgos principales.
* Automatización de la Generación de Informes:
  + Escribir un script en Python que genere informes a partir de los datos del dataset y las plantillas diseñadas.
  + Permitir la personalización de los informes según las necesidades específicas del cliente.

**Rol de la ciencia de datos**: La ciencia de datos jugará un papel crucial en la implementación de esta idea, permitiendo el procesamiento y análisis eficiente de grandes volúmenes de datos. Utilizando técnicas avanzadas de análisis y visualización de datos, DataVista Analytics podrá transformar datos brutos en insights accionables, facilitando la toma de decisiones informadas.

## Propuesta de valor

Los beneficios esperados al implementar los informes de mercado personalizados incluyen:

* Aumento de ingresos: Ofrecer un nuevo servicio genera una nueva fuente de ingresos para la empresa.
* Posicionamiento como experto: DataVista se consolidaría como un referente en análisis de datos para el sector automotriz.
* Mayor visibilidad: Atraer nuevos clientes y consolidar la relación con los existentes.
* Ventaja competitiva: Ofrecer un servicio innovador y de alto valor agregado.
* Expansión a nuevos mercados: Los informes pueden adaptarse a diferentes necesidades y mercados.
* Mejora de la toma de decisiones: Tanto para DataVista como para sus clientes.
* Fidelización de clientes: al ofrecer un servicio personalizado, DataVista podría fidelizar clientes y generar un flujo de ingresos constante.

La implementación de informes de mercado personalizados no solo resolverá problemas críticos para los clientes, sino que también ofrecerá múltiples beneficios estratégicos para DataVista Analytics, posicionándola como un líder en el análisis de datos en el mercado automotriz.

# Expectativas y Objetivos del Proyecto

## Objetivos a corto plazo:

**Semana 1: Preparación y análisis inicial**

* Revisión del dataset:
  + Revisar y limpiar el dataset df\_autos para asegurar la calidad de los datos.
  + Identificar las variables clave y posibles inconsistencias.
  + Entregable: Informe de calidad de datos y plan de limpieza.
* **Definición de requisitos:**
  + Reunirse con los stakeholders (concesionarios, fabricantes, empresas de seguros) para entender sus necesidades y expectativas.
  + Definir las preguntas clave que los informes deben responder.
  + Entregable: Documento de requisitos y expectativas del cliente.

**Semana 2: Desarrollo de herramientas de análisis**

* Desarrollo de funciones en python:
  + Escribir funciones en Python para procesar y analizar los datos del dataset.
  + Utilizar librerías como Pandas para manipulación de datos.
  + Entregable: Código funcional para el procesamiento de datos.
* Generación de visualizaciones:
  + Crear visualizaciones iniciales utilizando Matplotlib y Seaborn.
  + Generar gráficos y tablas que representen las tendencias y patrones identificados.
  + Entregable: Conjunto de visualizaciones preliminares.

**Semana 3: Diseño de plantillas de informes**

* Diseño de plantillas:
  + Diseñar plantillas de informes con una estructura clara y visualmente atractiva.
  + Incluir secciones para gráficos, tablas y resúmenes de hallazgos.
  + Entregable: Plantillas de informes en formato editable.
* Revisión y feedback:
  + Presentar las plantillas a los stakeholders para obtener feedback.
  + Ajustar el diseño según las sugerencias recibidas.
  + Entregable: Plantillas revisadas y aprobadas.

**Semana 4: Desarrollo de prototipos de informes**

* Creación de prototipos:
  + Generar prototipos de informes utilizando los datos del dataset y las plantillas diseñadas.
  + Incluir ejemplos de gráficos y análisis detallados.
  + Entregable: Prototipos de informes listos para revisión.
* Validación de prototipos:
  + Validar los prototipos con los stakeholders para asegurar que cumplen con los requisitos.
  + Realizar ajustes necesarios basados en el feedback.
  + Entregable: Prototipos validados y ajustados.

**Semana 5: Automatización de la generación de informes**

* Desarrollo del script de automatización:
  + Escribir un script en Python que automatice la generación de informes a partir de los datos del dataset y las plantillas diseñadas.
  + Incluir opciones de personalización según las necesidades específicas del cliente.
  + Entregable: Script de automatización funcional.
* Pruebas de automatización:
  + Realizar pruebas exhaustivas del script para asegurar su correcto funcionamiento.
  + Ajustar y optimizar el script según los resultados de las pruebas.
  + Entregable: Script probado y optimizado.

**Semana 6-7: Implementación y pruebas piloto**

* Implementación inicial:
  + Implementar el sistema de generación de informes en un entorno de prueba.
  + Realizar pruebas piloto con clientes seleccionados.
  + Entregable: Sistema implementado y pruebas piloto realizadas.
* Recopilación de feedback:
  + Recopilar feedback de los clientes piloto sobre la utilidad y precisión de los informes.
  + Realizar ajustes y mejoras basadas en el feedback recibido.
  + Entregable: Informe de feedback y plan de mejoras.

**Semana 8-11: Lanzamiento y comercialización**

* Lanzamiento del servicio:
  + Lanzar el servicio de informes de mercado personalizados al mercado.
  + Establecer relaciones con concesionarios, fabricantes y empresas de seguros para ofrecerles los informes.
  + Entregable: Servicio lanzado y disponible para clientes.
* Monitoreo y optimización:
  + Monitorear el uso y la efectividad de los informes.
  + Continuar optimizando el servicio basado en el feedback continuo de los clientes.
  + Entregable: Informes de monitoreo y optimización continua.

## Objetivos a largo plazo:

* Impacto en la empresa:
  + Posicionar a DataVista Analytics como líder en el análisis de datos dentro del sector automotriz.
  + Diversificar las fuentes de ingresos mediante la venta de informes personalizados.
  + Aumentar la fidelización y satisfacción de los clientes actuales.
* Impacto en el mercado:
  + Proveer a los actores del mercado automotriz con herramientas de análisis avanzadas para tomar decisiones informadas.
  + Facilitar la identificación de tendencias y oportunidades de mercado, mejorando la competitividad del sector.

## Objetivos del Equipo:

Fase 1: Análisis y Desarrollo (Semanas 1-3)

* Hito 1: Completar el análisis inicial del dataset df\_autos.
* Hito 2: Desarrollar las funciones de análisis en Python.
* Entregable: Código funcional para el procesamiento y análisis de datos.

Fase 2: Diseño y Prototipado (Semanas 4-5)

* Hito 3: Diseñar plantillas de informes con gráficos y tablas.
* Hito 4: Crear prototipos de informes personalizados.
* Entregable: Plantillas de informes y prototipos listos para revisión.

Fase 3: Automatización y Personalización (Semanas 6-7)

* Hito 5: Desarrollar el script de automatización para la generación de informes.
* Hito 6: Implementar opciones de personalización en los informes.
* Entregable: Script de automatización y opciones de personalización funcionales.

Fase 4: Implementación y Lanzamiento (Semanas 8-11)

* Hito 7: Realizar pruebas piloto con clientes seleccionados.
* Hito 8: Ajustar y optimizar los informes basados en feedback.
* Hito 9: Lanzar el servicio de informes de mercado personalizados.
* Entregable: Servicio de informes lanzado y disponible para clientes.

# Metas Individuales

## Roles y Responsabilidades

**Chief Data Officer (CDO):**

* **Descripción del rol:** La Coordinadora de la Tecnicatura Superior en Ciencia de Datos e Inteligencia Artificial será responsable de la visión integral del proyecto. Su función principal será asegurar que la solución esté alineada con los objetivos educativos y tecnológicos del equipo. Además, supervisará el progreso general, proporcionando orientación estratégica para garantizar el éxito del proyecto.
* **Tareas y responsabilidades:**
  + Monitorear y supervisar el progreso del equipo.
  + Asegurar que la solución propuesta esté alineada con los objetivos educativos y tecnológicos.
  + Proveer retroalimentación estratégica y garantizar la calidad del trabajo final.

**Supervisor:**

* **Descripción del rol:** El profesor del curso servirá como supervisor, guiando al equipo en la correcta implementación del proyecto. Será responsable de asegurar que los métodos utilizados sean académicamente válidos y que se mantenga la calidad en el desarrollo del modelo predictivo y la documentación.
* **Tareas y responsabilidades:**
  + Proveer retroalimentación en cada fase del proyecto.
  + Monitorear y asegurar la calidad técnica y documental de los entregables.

**Responsable de Equipo de Trabajo:**

* **Descripción del rol:** Este rol será rotativo entre los miembros del equipo y garantizará que todas las tareas se completen a tiempo y que la comunicación entre los roles del proyecto fluya correctamente.
* **Tareas y responsabilidades:**
  + Coordinar las actividades diarias del equipo.
  + Asegurar que los hitos del proyecto se cumplan según el cronograma establecido.
  + Gestionar la comunicación entre el equipo y los supervisores.

**Metas individuales:**

* Mantener un flujo de trabajo eficiente y organizado.
* Asegurar la entrega de los hitos del proyecto en el tiempo estimado.

**Plan de desarrollo personal:**

* **Desarrollo de liderazgo:** Mejorar las habilidades de liderazgo en la gestión de equipos multifuncionales, aprendiendo a delegar tareas de manera efectiva y asegurar una comunicación clara y precisa.
* **Gestión de proyectos:** Aumentar la capacidad de seguimiento y control de múltiples fases del proyecto, dominando herramientas de gestión de proyectos y optimizando la coordinación del equipo para asegurar entregables de alta calidad.

**Primer Reemplazo:**

* **Descripción del rol:** Asumirá las responsabilidades del Responsable de Equipo en su ausencia, garantizando la continuidad del proyecto.
* **Tareas y responsabilidades:**
  + Tomar decisiones operativas en ausencia del Responsable de Equipo.
  + Mantener al equipo informado sobre el progreso del proyecto.

**Metas individuales:**

* Asegurar la continuidad en la gestión y el progreso del proyecto.
* Mantener una vista general clara de todas las tareas en curso.

**Plan de desarrollo personal:**

* **Adaptabilidad y toma de decisiones:** Fortalecer las habilidades de toma de decisiones rápidas en situaciones imprevistas, mejorando la capacidad de adaptación frente a cambios en el entorno de trabajo.
* **Liderazgo de emergencia:** Desarrollar habilidades de liderazgo cuando se enfrenta a situaciones de presión, siendo capaz de asumir roles de responsabilidad cuando sea necesario, y optimizar la gestión del tiempo para cumplir con los plazos.

**Segundo Reemplazo:**

* **Descripción del rol:** Actuará cuando tanto el Responsable de Equipo como el Primer Reemplazo estén ausentes, asegurando que el proyecto continúe sin interrupciones.
* **Tareas y responsabilidades:**
  + Supervisar el progreso del equipo en caso de emergencia.
  + Mantener la coordinación entre los miembros del equipo para garantizar que se cumplan los plazos del proyecto.

**Metas individuales:**

* Estar siempre preparado para asumir responsabilidades en cualquier momento.
* Mantener una supervisión continua del progreso del proyecto.

**Plan de desarrollo personal:**

* **Preparación y proactividad:** Desarrollar una mentalidad proactiva, anticipando posibles problemas y manteniéndose siempre informado sobre el progreso del proyecto.
* **Gestión del estrés y resolución de problemas:** Mejorar la capacidad para gestionar el estrés en situaciones críticas, fortaleciendo la resolución de problemas y la capacidad para asumir decisiones en momentos de alta presión.

# Plan de Acción

## Fases del Proyecto

**Fase 1: Análisis Inicial y Preparación (Semanas 1-2)**

* Actividades:
  + Revisión del dataset: Identificar posibles inconsistencias, valores atípicos y datos faltantes en el dataset df\_autos.
  + Limpieza de datos: Implementar técnicas de limpieza como imputación de datos faltantes, eliminación de duplicados y estandarización de variables.
  + Identificación de variables clave: Determinar las variables que serán esenciales para el análisis y la creación de informes personalizados.
* Recursos necesarios: Dataset df\_autos, librerías de Python (Pandas, NumPy), equipo de analistas de datos.
* Plazos: 2 semanas.

**Fase 2: Desarrollo de Herramientas de Análisis (Semanas 3-4)**

* Actividades:
  + Desarrollo de funciones en Python: Creación de funciones específicas para el análisis exploratorio de datos, incluyendo estadísticas descriptivas y visualizaciones.
  + Generación de visualizaciones avanzadas: Uso de Matplotlib y Seaborn para crear gráficos que muestren tendencias, distribuciones y correlaciones clave.
  + Definición de métricas de éxito: Establecer métricas para evaluar la efectividad de los análisis y la calidad de los informes generados.
* Recursos necesarios: Python, Matplotlib, Seaborn, Jupyter Notebooks, equipo técnico.
* Plazos: 2 semanas.

Fase 3: Diseño y Prototipado de Informes (Semanas 5-7)

* Actividades:
  + Diseño de plantillas de informes: Crear plantillas estandarizadas que incluyan gráficos, tablas y resúmenes, adaptadas a diferentes tipos de clientes.
  + Prototipado de informes: Generación de prototipos de informes utilizando datos reales del dataset para validar el diseño y el contenido.
  + Revisión con stakeholders: Presentación de prototipos a los clientes potenciales para recibir feedback y realizar ajustes necesarios.
* Recursos necesarios: Herramientas de diseño gráfico, software de edición de documentos (LaTeX, MS Word), equipo de diseño y marketing.
* Plazos: 3 semanas.

Fase 4: Automatización y Desarrollo de Personalización (Semanas 8-9)

* Actividades:
  + Desarrollo de scripts de automatización: Programación de scripts en Python para automatizar la generación de informes a partir del dataset y las plantillas diseñadas.
  + Personalización de informes: Implementación de opciones de personalización para que los informes se adapten a las necesidades específicas de cada cliente.
  + Pruebas de funcionalidad: Realización de pruebas exhaustivas para asegurar que el proceso automatizado funcione sin errores y cumpla con los requisitos.
* Recursos necesarios: Python, librerías de automatización (Jinja2, Pandas), servidores de prueba, equipo técnico.
* Plazos: 2 semanas.

Fase 5: Implementación, Pruebas Piloto y Lanzamiento (Semanas 10-12)

* Actividades:
  + Implementación en entorno de producción: Despliegue del sistema de generación de informes en un entorno real.
  + Pruebas piloto: Colaboración con un grupo selecto de clientes para realizar pruebas piloto, recopilando datos de rendimiento y feedback.
  + Lanzamiento comercial: Realización del lanzamiento oficial del servicio, incluyendo la creación de campañas de marketing y estrategias de venta.
* Recursos necesarios: Infraestructura de servidores, herramientas de monitorización (Prometheus, Grafana), equipo de ventas y marketing.
* Plazos: 3 semanas.

## Hitos y Entregables

* Hito 1: Finalización del análisis inicial y limpieza de datos (Semana 2)
  + Entregable: Informe de calidad de datos y plan de limpieza detallado.
  + Fechas de revisión: 29/08.
* Hito 2: Desarrollo y validación de funciones de análisis (Semana 4)
  + Entregable: Código Python para análisis y visualizaciones, conjunto de gráficos y tablas preliminares.
  + Fechas de revisión: 10/09.
* Hito 3: Diseño y validación de plantillas de informes (Semana 7)
  + Entregable: Plantillas de informes aprobadas por stakeholders y prototipos de informes.
  + Fechas de revisión: 28/09.
* Hito 4: Desarrollo del script de automatización y personalización (Semana 9)
  + Entregable: Script funcional de automatización con opciones de personalización.
  + Fechas de revisión: 10/10.
* Hito 5: Lanzamiento del servicio de informes personalizados (Semana 12)
  + Entregable: Servicio operativo, informes generados y entregados a clientes.
  + Fechas de revisión: 04/11.

## Metodología de Trabajo

Frecuencia de Reuniones:

* Reuniones iniciales: Reuniones diarias durante las primeras dos semanas para asegurar una comprensión clara del proyecto y alineación de objetivos.
* Reuniones semanales: A partir de la tercera semana, reuniones semanales para revisar el progreso, resolver problemas y ajustar el plan según sea necesario.
* Reuniones de revisión: Reuniones al final de cada fase para evaluar los entregables y preparar la siguiente etapa.

Herramientas de Comunicación y Colaboración:

* Google Drive: Para la gestión de documentos compartidos y colaboración en tiempo real.
* Trello: Para el seguimiento de tareas y gestión de proyectos, incluyendo la asignación de tareas y la monitorización del progreso.
* GitHub: Para el control de versiones del código y la colaboración en el desarrollo de scripts.

Proceso de Seguimiento y Revisión:

* Monitoreo del progreso: Uso de herramientas como Trello para seguir el avance de cada tarea, asignando responsables y fechas límite.
* Revisión continua: Evaluaciones periódicas de los entregables por parte del equipo y los stakeholders para asegurar que se cumplen los requisitos y se mantienen los estándares de calidad.
* Gestión de riesgos: Identificación temprana de posibles obstáculos o problemas, con planes de contingencia preparados para asegurar la continuidad del proyecto.