

Tecnicatura en Ciencia de Datos e Inteligencia Artificial  
Módulo Analista de Datos

# Análisis de datos de Recursos Humanos de la empresa BioTech Innovations S.A.

## EQUIPO DOCENTE

Ugarte, Marcos  
Pratta, Nahuel

## GRUPO DATAMINDS

1. Angiolini, Sofía Carla
2. Ayán, María Trinidad
3. Donati, Camila
4. Giordano, Ariel Eduardo
5. Herrera, Edgar Fabián
6. López, María Gabriela
7. Quiroga, Fernanda
8. Valvassori, Etelvina Graciela



## Índice

Introducción	2
Objetivos	2
Desarrollo	3
Encuesta de satisfacción y condiciones laborales en BioTech Innovations S.A.	3
Descripción del dataset y de los datos de cada columna	7
Metodología de trabajo	10
Dejamos asentado en este documento que la compañera María Gabriela López ha abandonado el proyecto por motivos personales.	14
Conclusión general	14
Bibliografía	15

## **Introducción**

BioTech Innovations S.A. es una empresa líder en el campo de la biotecnología y las ciencias de la vida. Fundada hace 15 años, la empresa se dedica a la investigación, desarrollo y comercialización de soluciones innovadoras para la industria médica y científica. Con un enfoque en la mejora de la calidad de vida a través de la innovación tecnológica, BioTech Innovations combina la investigación de vanguardia con la aplicación práctica para ofrecer productos y servicios de alta calidad.

BioTech Innovations S.A. se esfuerza por mantener un ambiente de trabajo dinámico y satisfactorio, apoyando a sus empleados en su desarrollo profesional y personal. Con un compromiso constante hacia la excelencia y la innovación, la empresa sigue creciendo y marcando la diferencia en el campo de la biotecnología.

La empresa considera a los empleados como el recurso más valioso. A continuación, se presenta un estudio realizado por el área de recursos humanos de la empresa, que desea estudiar los perfiles representativos de los empleados, destacando su diversidad en términos de edad, roles laborales y niveles de satisfacción.

## **Objetivos**

### **Objetivos generales del proyecto:**

- Realizar un análisis completo de un conjunto de datos sin depurar.
- Utilizar herramientas gráficas, medidas de posición y dispersión para obtener conclusiones sobre los datos.
- Aplicar los conocimientos adquiridos en el módulo Analista de Datos.

## **Objetivos específicos del proyecto:**

- Seleccionar y limpiar adecuadamente el conjunto de datos a analizar.
- Aplicar técnicas estadísticas y exploratorias para comprender las características y relaciones entre las variables presentes en los datos.
- Utilizar gráficos para visualizar patrones o tendencias en los datos.
- Calcular medidas descriptivas como media, mediana, desviación estándar, etc. para describir la distribución y variabilidad en los datos.
- Extraer conclusiones relevantes basadas en el análisis realizado.

## **Desarrollo**

### **Encuesta de satisfacción y condiciones laborales en BioTech Innovations S.A.**

Estimado empleado/a,

En BioTech Innovations S.A. valoramos tu opinión y estamos comprometidos en mejorar continuamente el entorno laboral. Te pedimos que completes esta encuesta para ayudarnos a entender mejor tus experiencias y necesidades. La información proporcionada será confidencial y se utilizará únicamente para mejorar nuestras políticas y prácticas laborales.

### **Instrucciones para completar esta encuesta:**

- Por favor, completa todas las preguntas con la mayor precisión posible.
- Marca las opciones que correspondan.
- Una vez completada la encuesta, deposítala en la caja de sugerencias anónimas o envíala al Departamento de Recursos Humanos.

## 1. Información Demográfica

Edad: \_\_\_\_\_

Género:

☐ Masculino

☐ Femenino

Estado civil:

☐ Soltero

☐ Casado

☐ Separado

☐ Divorciado

☐ Viudo

## 2. Datos de Empleo

Renunció:

☐ Sí

☐ No

Departamento:

☐ Ventas

☐ Investigación y Desarrollo

☐ Recursos humanos

Rol laboral:

☐ Ejecutivo de Ventas

☐ Investigador Científico

- ☐ Técnico de Laboratorio
- ☐ Director de Manufactura
- ☐ Representante de Salud
- ☐ Gerente
- ☐ Representante de Ventas
- ☐ Director de Investigación
- ☐ Recursos Humanos

Distancia desde casa al trabajo (en km.): \_\_\_\_\_

Años en el rol actual: \_\_\_\_\_

Años desde el último ascenso: \_\_\_\_\_

Años con el gerente actual: \_\_\_\_\_

Antigüedad en "BioTech Innovations S.A." : \_\_\_\_\_

### **3. Educación y Capacitación**

Nivel de educación:

- ☐ 1 (Secundario)
- ☐ 2 (Técnico)
- ☐ 3 (Universitario)
- ☐ 4 (Maestría)
- ☐ 5 (Doctorado)

Campo de estudio:

- ☐ Ciencias de la Vida
- ☐ Medicina
- ☐ Marketing
- ☐ Recursos Humanos

☐ Técnico☐ Otro

Número de capacitaciones recibidas el año pasado: \_\_\_\_\_

#### **4. Satisfacción y Ambiente Laboral**

¿Con qué frecuencia viajas por motivos de trabajo?

☐ Rara vez viaja☐ Viaja frecuentemente☐ No viaja

¿Cómo calificarías en general tu satisfacción con el entorno laboral?

☐ 1 (Muy insatisfecho/a)☐ 2 (Poco satisfecho)☐ 3 (Satisfecho)☐ 4 (Muy satisfecho)

¿Cómo calificarías tu satisfacción laboral general?

☐ 1 (Muy insatisfecho/a)☐ 2 (Poco satisfecho)☐ 3 (Satisfecho)☐ 4 (Muy satisfecho)

¿Cómo calificarías tu satisfacción con las relaciones laborales?

☐ 1 (Muy insatisfecho/a)☐ 2 (Poco satisfecho)☐ 3 (Satisfecho)

☐ 4 (Muy satisfecho)

¿Cuál fue el porcentaje de incremento salarial que recibiste en el último año?

\_\_\_\_\_

¿Cuál es tu jornada laboral estándar (en horas)? \_\_\_\_\_

**¡Gracias por tu participación y por ayudarnos a hacer de BioTech Innovations  
S.A. un mejor lugar para trabajar!**

### Descripción del dataset y de los datos de cada columna

Debido a la dificultad que encontramos al elegir un dataset que cumpliera con los requisitos, decidimos crear uno nosotros mismos para lo cual combinamos un excel de RRHH y lo fuimos modificando basándonos en una encuesta realizada por RRHH y reportada en el libro de Berenson, M., & Levine, D. (1996). Dado que el dataset ya estaba depurado, tuvimos que incorporar algunos datos erróneos para poder realizar el análisis requerido por el proyecto.

Nombre	Tipo de dato	Valores que puede tomar
Edad	Int	18 a 70
Renunció	Strings	- Si - No
Viaja por trabajo	Strings	- Viaja frecuentemente - Rara vez viaja - No viaja



Nombre	Tipo de dato	Valores que puede tomar
Departamento	Strings	<ul style="list-style-type: none"> <li>- Ventas</li> <li>- Investigación y Desarrollo</li> <li>- Recursos Humanos</li> </ul>
Distancia a su casa (km)	Int	0 a 100
Nivel de educación	Int	1 a 5
Campo de estudio	Strings	<ul style="list-style-type: none"> <li>- Ciencias de la Vida</li> <li>- Medicina</li> <li>- Marketing</li> <li>- Recursos Humanos</li> <li>- Técnico</li> <li>- Otro</li> </ul>
Número de empleado	Int	1 a 5000
Entorno laboral	Int	1 a 4
Género	Strings	<ul style="list-style-type: none"> <li>- Femenino</li> <li>- Masculino</li> </ul>
Rol Laboral	Strings	<ul style="list-style-type: none"> <li>- Ejecutivo de Ventas</li> <li>- Investigador Científico</li> <li>- Técnico de Laboratorio</li> <li>- Director de Manufactura</li> <li>- Representante de Salud</li> <li>- Gerente</li> </ul>

Nombre	Tipo de dato	Valores que puede tomar
		<ul style="list-style-type: none"> <li>- Representante de Ventas</li> <li>- Director de Investigación</li> <li>- Recursos Humanos</li> </ul>
Satisfacción laboral general	Int	1 a 4
Estado civil	Strings	<ul style="list-style-type: none"> <li>- Soltero</li> <li>- Casado</li> <li>- Separado</li> <li>- Divorciado</li> <li>- Viudo</li> </ul>
Porcentaje Aumento Del Salario	Int	0 a 100
Satisfacción de relaciones	Int	1 a 4
Cantidad de horas trabajadas	Int	80 a 160
Antigüedad laboral	Int	0 a 55
Antigüedad en la empresa	Int	0 a 55
Capacitaciones del año anterior	Int	0 a 10
Antigüedad en el mismo puesto	Int	0 a 55

Nombre	Tipo de dato	Valores que puede tomar
Años desde el último ascenso	Int	0 a 55
Años con el mismo gerente	Int	0 a 55

## Metodología de trabajo

Para desarrollar el presente trabajo, tomamos las herramientas del marco de trabajo denominado Scrum. Esta metodología toma su nombre de una de las formaciones del rugby, en la cual los jugadores tienen que sacar la pelota sin tocarla con las manos. Para lograr esta formación, el equipo debe tener una gran coordinación y cooperación, por lo que Scrum, como gestión de proyectos, se puede definir como un modelo de desarrollo ágil y flexible.

El primer paso para desarrollar las consignas planteadas por las cátedras fue la creación de un grupo de WhatsApp denominado “DataMinds ISPC - Grupo Proyecto” con los integrantes del equipo de trabajo. Una vez creado el grupo informativo, elegimos a quién iba a desempeñar el rol de Scrum Máster. El Scrum Máster facilita los eventos Scrum según se necesite y se asegura de que los colaboradores adhieran correctamente a las recomendaciones de la guía de trabajo Scrum.

✕ Info. del grupo



Captura de pantalla del grupo de trabajo en WhatsApp

Los eventos SCRUM son la reunión de planificación del Sprint, el Scrum diario, el Sprint, la revisión del Sprint y la retrospectiva. Cada Sprint representa una Evidencia de Aprendizaje, por lo que el equipo de trabajo DataMinds ha logrado tres Sprints:

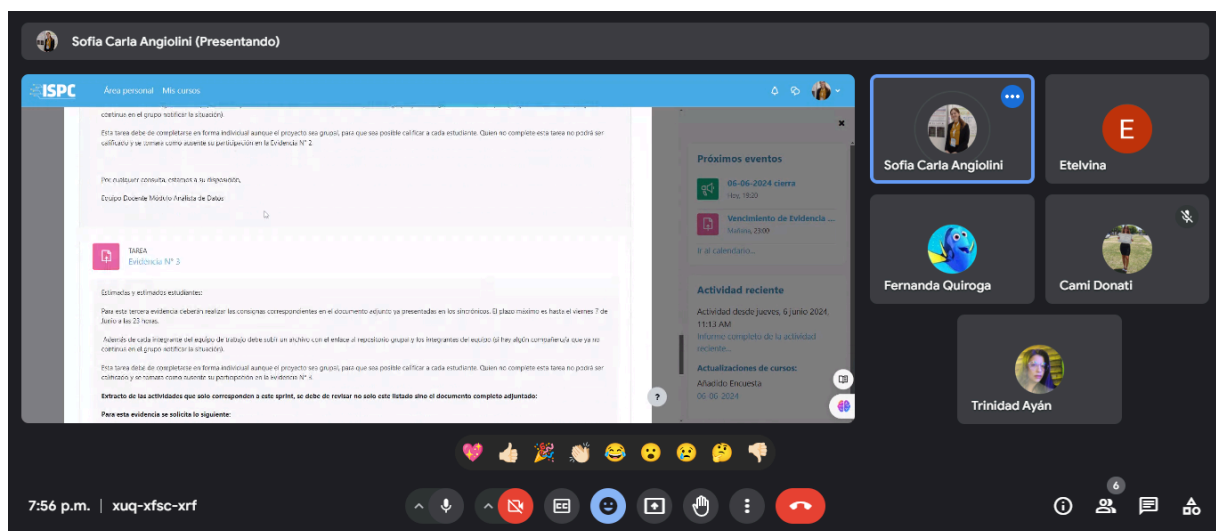
- Sprint 1: conformación del equipo de trabajo y creación del repositorio en GitHub.
- Sprint 2: obtener un dataset con los datos sin depurar para analizar las columnas numéricas y categóricas; analizar la cantidad de valores válidos sobre el total en las variables numéricas; realizar gráficos; obtener las

medidas de tendencia central y variación de las variables numéricas; y, por último, los valores ausentes y mal registrados o erróneos.

- Sprint 3: usar la librería Pandas para la limpieza de los datos y usar la librería Matplotlib para realizar gráficos que guarden relación con los datos, tanto en variables categóricas como numéricas.

Antes de cada Sprint, el grupo realiza una reunión para planificar el Sprint y revisar los requerimientos a cumplir (backlog). Estas reuniones mediante la plataforma de videollamada Meet fueron tres:

- Reunión de backlog 1: jueves 2 de mayo
- Reunión de backlog 2: sábado 18 de mayo
- Reunión de backlog 3: lunes 3 de junio



Captura de pantalla de una reunión por la plataforma Meet

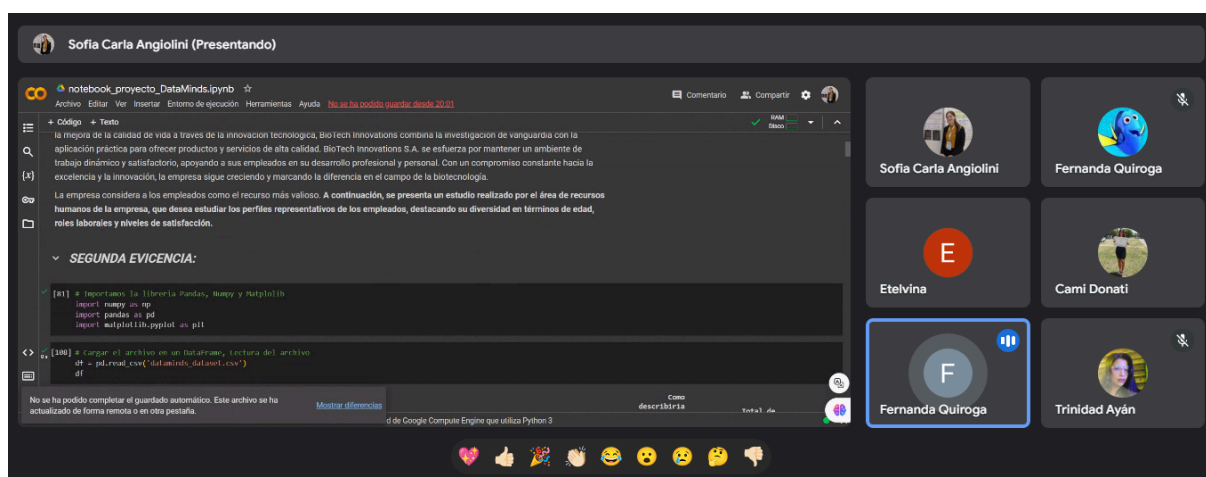
El Scrum diario es otra de las ceremonias del enfoque Scrum. En este caso, por cuestiones de disponibilidad horaria de los integrantes, DataMinds no realizó

encuentros diarios. Sin embargo, el grupo de WhatsApp antes mencionado sirvió como medio de contacto para comunicar los avances obtenidos del Sprint. Las reuniones para realizar el Sprint fueron las siguientes:

- Sábado 4 de mayo
- Martes 21 de mayo
- Miércoles 7 de junio

Por último, cada Sprint tuvo una reunión con el objetivo de realizar la revisión del Sprint y revisar en conjunto si la Evidencia de Aprendizaje terminada cumplía con los requisitos pedidos en cada trabajo. Estas instancias siempre se realizaron antes de adjuntar la entrega:

- Lunes 6 de mayo
- Viernes 24 de mayo
- Viernes 7 de junio



Captura de pantalla de una reunión por la plataforma Meet

Dejamos asentado en este documento que la compañera María Gabriela López ha abandonado el proyecto por motivos personales.

## **Conclusión general**

El presente trabajo se ha centrado en el análisis exhaustivo de datos sobre la satisfacción y condiciones laborales en BioTech Innovations S.A. con el objetivo de extraer conclusiones relevantes y proporcionar recomendaciones para mejorar el entorno laboral. Se implementaron diversas técnicas y metodologías, desde la limpieza y depuración de datos hasta la visualización gráfica y el análisis estadístico.

El análisis de los datos permitió identificar patrones y tendencias en variables como la antigüedad de los empleados, la satisfacción laboral y las condiciones de trabajo. Utilizando herramientas como Pandas y Matplotlib, se realizaron operaciones de limpieza y transformación de datos, asegurando la calidad y precisión del análisis final. Este análisis inicial permite una visión clara del estado y calidad del dataset, posibilitando análisis más detallados como correlaciones entre variables, por ejemplo: edad y satisfacción laboral; segmentación de empleados por departamentos y análisis de su satisfacción; e identificación de patrones en las renuncias y su relación con otras variables. Este proceso garantiza que los datos estén limpios y listos para análisis avanzados que pueden proporcionar *insights* valiosos para BioTech Innovations S.A.

Asimismo, este estudio proporciona una base sólida para implementar estrategias que mejoren la retención y satisfacción de los empleados, subrayando la importancia de políticas de desarrollo profesional y la consideración de la diversidad de perfiles en la empresa.

La adopción de la metodología Scrum facilitó una colaboración efectiva entre los miembros del equipo, asegurando una comunicación fluida y una gestión ágil del

proyecto. Cada sprint permitió alcanzar objetivos específicos y evaluar continuamente los avances, garantizando la calidad y coherencia del trabajo realizado.

En resumen, este proyecto no solo cumplió con los objetivos planteados, sino que también proporcionó un marco de referencia valioso para futuros análisis y estudios. El uso de herramientas de análisis de datos y la aplicación de metodologías ágiles demostraron ser efectivas para realizar un estudio detallado y significativo.

## **Bibliografía**

- ❖ Berenson, M., & Levine, D. (1996). *Estadística básica en administración: Conceptos y aplicaciones* (6ª ed.). Prentice Hall Hispanoamericana.
- ❖ Mayer, T. (2014). *Por un Scrum popular: notas para una revolución ágil*. Dymaxicon.
- ❖ Triola, M. (2008). *Estadística* (10ª ed.). Pearson Addison Wesley.