



UNIVERSIDAD DE  
SAN BUENAVENTURA  
BOGOTÁ

# Taller Minería descriptiva y minería predictiva

Proyecto integrador

Jorge Andrés Ballen

David Alejandro Jáuregui Merchán

Dayana Andrea Escobar Medina

Brayan Steven López

Universidad de San Buenaventura

Facultad de ingeniería

Tecnología en desarrollo de software

Minería de datos

Noviembre 2025



## Contenido

1. Instrucciones del taller .....	3
2. Desarrollo del taller .....	3
2.1 Definición del problema y objetivos .....	3
2.2 Selección y preparación de dataset .....	3
<b>Selección de datos</b> .....	4
<b>Preparación de datos</b> .....	5
2.3 Aplicación de técnicas descriptivas y predictivas .....	6
<b>Técnica descriptiva</b> .....	6
<b>Técnica predictiva</b> .....	8



## 1. Instrucciones del taller

- A. Definición del problema y objetivos.
- B. Selección y preparación de dataset.
- C. Aplicación de técnicas descriptivas y predictivas.
- D. Evaluación y presentación de resultados.

## 2. Desarrollo del taller

### 2.1 Definición del problema y objetivos

**Objetivo general:** Aplicar técnicas de minería de datos descriptiva y predictiva para analizar los hábitos financieros de los encuestados y determinar los factores que influyen en la adopción de herramientas de control de presupuesto.

**Problema:** La mayoría de las personas no aplican métodos ni herramientas para controlar su presupuesto personal, lo que puede afectar su estabilidad financiera y capacidad de ahorro.

#### Objetivos específicos:

- Identificar el nivel de conocimiento sobre la importancia del presupuesto.
- Analizar las causas que impiden mantener un hábito financiero constante.
- Evaluar la relación entre conocimiento financiero y uso de herramientas.
- Proponer recomendaciones para el diseño de una aplicación o sistema que fomente el control presupuestal.

### 2.2 Selección y preparación de dataset

Los datos por usar en nuestra dataset se originan en una encuesta de caracterización realizada por el equipo de desarrollo del proyecto Boss Budget.



## Selección de datos

- Importamos los datos que se descargaron desde Microsoft Forms en formato xlsx (Excel).

The screenshot shows the 'Navegador' (Navigator) dialog box in Microsoft Excel. On the left, the 'Sheet1' tab is selected in the list of tables. The main area displays a table with 14 rows of data, each containing an ID, start time, end time, and email address. The table is titled 'Sheet1' and has columns for 'ID', 'Hora de inicio', 'Hora de finalización', and 'Correo electrónico'. The data shows various times from April to October 2023. A note at the bottom indicates that the preview data is truncated. On the right, the ribbon shows tabs for 'Publicar', 'Preparar', 'Compartir', and 'Datos'. The 'Datos' tab is selected, showing a 'Buscar' (Search) field and a 'Obtener' (Get) button.

ID	Hora de inicio	Hora de finalización	Correo electrónico
1	4/10/2023 6:58:29 p. m.	4/10/2023 7:30:10 p. m.	anonymous
2	4/10/2023 8:09:52 p. m.	4/10/2023 8:17:47 p. m.	anonymous
3	6/10/2023 11:50:26 a. m.	6/10/2023 11:53:06 a. m.	anonymous
4	6/10/2023 12:16:25 p. m.	6/10/2023 12:18:02 p. m.	anonymous
5	6/10/2023 3:17:38 p. m.	6/10/2023 3:28:40 p. m.	anonymous
6	6/10/2023 10:14:20 p. m.	6/10/2023 10:16:11 p. m.	anonymous
7	6/10/2023 10:15:57 p. m.	6/10/2023 11:15:54 p. m.	anonymous
8	10/10/2023 4:07:39 p. m.	10/10/2023 4:10:32 p. m.	anonymous
9	10/10/2023 4:08:22 p. m.	10/10/2023 4:19:06 p. m.	anonymous
10	10/10/2023 4:33:25 p. m.	10/10/2023 4:35:43 p. m.	anonymous
11	11/10/2023 10:31:34 a. m.	11/10/2023 10:37:02 a. m.	anonymous
12	17/10/2023 1:27:36 p. m.	17/10/2023 1:35:59 p. m.	anonymous
13	17/10/2023 1:37:23 p. m.	17/10/2023 1:44:45 p. m.	anonymous
14	17/10/2023 5:03:00 p. m.	17/10/2023 5:06:00 p. m.	anonymous



## Preparación de datos

- Borramos valores que no se usarán ya que son nulls segun los requerimientos de la tabla, como lo es correo electronico (está en anonimo), nombre, total de puntos, comentarios del cuestionario, hora de ultima edición, y los puntos y comentarios de cada pregunta los cuales están vacíos.

Configuración de la consulta

**PROPIEDADES**

- Nombre: Sheet1
- Todas las propiedades

**PASOS APLICADOS**

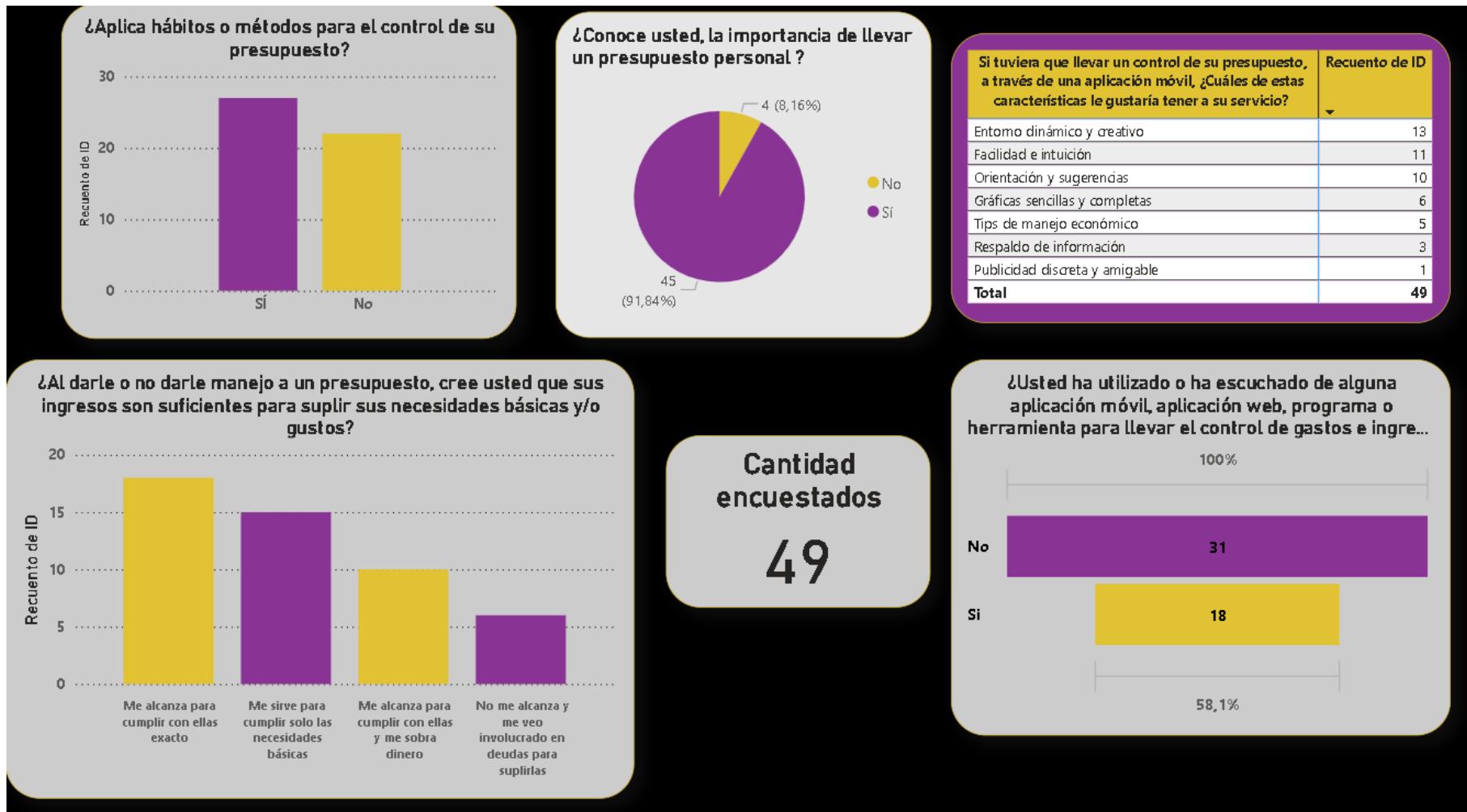
- Origen
- Navegación
- Encabezados promovidos
- Tipo cambiado
- Columnas quitadas

ID	Hora de inicio	Hora de finalización	Correo electrónico	¿Conoce usted, la impo
1	4/10/2023 6:58:29 p. m.	4/10/2023 7:30:10 p. m.	anonymous	Sí
2	4/10/2023 8:09:52 p. m.	4/10/2023 8:17:47 p. m.	anonymous	Sí
3	6/10/2023 11:50:26 a. m.	6/10/2023 11:53:06 a. m.	anonymous	Sí
4	6/10/2023 12:16:25 p. m.	6/10/2023 12:18:02 p. m.	anonymous	No
5	6/10/2023 3:17:38 p. m.	6/10/2023 3:28:40 p. m.	anonymous	Sí
6	6/10/2023 10:14:20 p. m.	6/10/2023 10:16:11 p. m.	anonymous	Sí
7	6/10/2023 10:15:57 p. m.	6/10/2023 11:15:54 p. m.	anonymous	Sí
8	10/10/2023 4:07:39 p. m.	10/10/2023 4:10:32 p. m.	anonymous	Sí
9	10/10/2023 4:08:22 p. m.	10/10/2023 4:19:06 p. m.	anonymous	Sí
10	10/10/2023 4:33:25 p. m.	10/10/2023 4:35:43 p. m.	anonymous	No
11	11/10/2023 10:31:34 a. m.	11/10/2023 10:37:02 a. m.	anonymous	Sí
12	17/10/2023 1:27:36 p. m.	17/10/2023 1:35:59 p. m.	anonymous	Sí
13	17/10/2023 1:37:23 p. m.	17/10/2023 1:44:45 p. m.	anonymous	Sí
14	17/10/2023 5:03:00 p. m.	17/10/2023 5:06:00 p. m.	anonymous	Sí



## 2.3 Aplicación de técnicas descriptivas y predictivas.

### Técnica descriptiva





**1. Aplicación de Métodos (Gráfico de Barras):**

A pesar de la alta conciencia, solo el 59.18% (29 de 49) de los encuestados afirma aplicar métodos o hábitos para el control de su presupuesto. Un significativo 40.82% (20 de 49) no lo hace, lo que indica una brecha entre el conocimiento y la acción.

**2. Importancia del Presupuesto (Gráfico de Torta):**

Existe una sólida mayoría que conoce la importancia de llevar un presupuesto personal, representando un 91.84% (45 de 49). Solo una pequeña fracción del 8.16% (4 personas) desconoce su importancia.

**3. Características Deseadas en un Servicio de Control (Tabla):**

Al considerar el control a través de una aplicación móvil/web, la característica más solicitada es un "Entorno dinámico y creativo" (13 votos), seguida de cerca por la "Realidad e intuición" (11 votos). Esto sugiere que los usuarios valoran una interfaz amigable y atractiva que facilite el proceso, más que solo las características técnicas como el "Respaldo de información" (3) o las "Gráficas sencillas y completas" (6).

**4. Suficiencia de Ingresos vs. Manejo Presupuestario (Gráfico de Barras):**

La mayoría, un 34.69% (17 encuestados), indica que sus ingresos les alcanzan para cumplir solo con las necesidades básicas y un 36.73% (18 encuestados) siente que sus ingresos les alcanzan exacto para suplir necesidades y gustos.

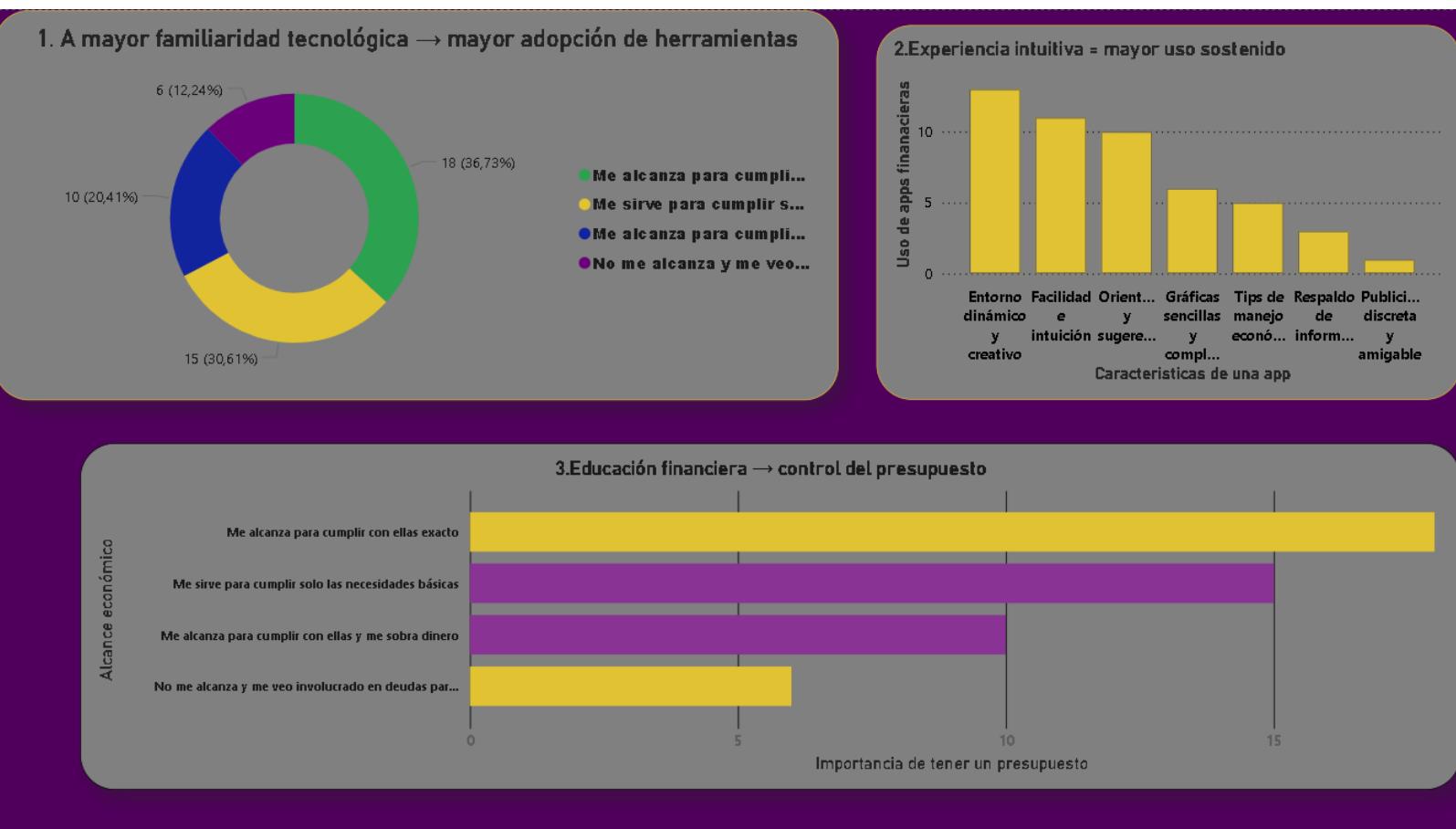
Un preocupante 12.24% (6 encuestados) afirma que no les alcanza y se ven involucrados en deudas para suplir sus necesidades, mientras que un 20.41% (10 encuestados) les sobra dinero, lo que se correlaciona con la aplicación de métodos esto nos indica que la mayoría de los encuestados (alrededor del 71%) se encuentra en una situación financiera ajustada ("Exacto" o "Solo básicas"), lo que subraya la urgencia de implementar un control presupuestario.

**5. Uso de Herramientas Digitales (Gráfico de Barras y Porcentaje):**

Una mayoría clara del 58.1% (31 encuestados) afirma no haber utilizado o escuchado de alguna aplicación o herramienta para el control de gastos/ingresos y solo el 41.9% (18 encuestados) ha utilizado o conoce estas herramientas. Esto confirma la oportunidad de mercado para una nueva herramienta digital, especialmente si se diseña con un "Entorno dinámico y creativo" como se solicitó, para capturar a la gran parte de la población que aún no usa estas soluciones.



## Técnica predictiva





### **1. Familiaridad Tecnológica Mayor Adopción (Gráfico Superior Izquierdo)**

Se predice un aumento en la adopción de herramientas digitales por parte de los segmentos con mayor necesidad (los que tienen ingresos ajustados o deudas). Estos grupos ( $36.73\% + 12.24\% = 48.97\%$ ) son los clientes potenciales más urgentes. La familiaridad tecnológica general seguirá creciendo, actuando como un multiplicador de adopción. Si la herramienta es fácil de usar, los usuarios con problemas de liquidez la adoptarán rápidamente como una solución de supervivencia financiera, no solo como una herramienta de optimización. La predicción es que el mercado pasará del estado actual de 42% de adopción a uno de 60% en los próximos 2-3 años, impulsado por la presión económica de los usuarios ajustados.

### **2. Experiencia Intuitiva ----> Mayor Uso Sostenido (Gráfico Superior Derecho)**

Las aplicaciones que inviertan en Gamificación, diseño UX/UI superior y automatización inteligente (Inteligencia Artificial para sugerencias personalizadas) dominarán el mercado. Las herramientas que se limiten a ser "libretas digitales" para registrar gastos tendrán una tasa de abandono alta., ya que el futuro del control presupuestario es la experiencia que se predice que la satisfacción del usuario (y por lo tanto, el uso sostenido) estará directamente correlacionada con el grado de "dinamismo creativo" e "intuición" que ofrezca la aplicación. Un usuario dejará de usar una app aburrida o compleja, incluso si le trae beneficios financieros.

### **3. Educación Financiera ---> Control del Presupuesto (Gráfico Inferior)**

La conciencia ya no es el cuello de botella; la ejecución lo es. Se predice un cambio de enfoque en el marketing de herramientas: dejará de enfatizar la "importancia" (que ya se conoce) y se centrará en la "facilidad de ejecución" y el "retorno de la inversión" (el dinero que el usuario puede ahorrar). El segmento en deuda ("No me alcanza y me veo involucrado en deudas"), aunque percibe la importancia (5 de 6), requiere un soporte más proactivo y "de rescate". Las herramientas futuras deberán ofrecer funciones de alerta de deuda y planes de pago automatizados, pasando de ser una herramienta de registro a un asesor financiero digital.