



Proyecto de Investigación Capstone

“Tucan S.A.”

Trabajo de Grado

Daniela Tole Silva

Isabella Ávila Ávila

Carlos Alberto Quintero Sánchez

Bogotá D.C

2023



Proyecto de Investigación Capstone

“Tucan S.A.”

Trabajo de Grado

Daniela Tole Silva

Isabella Ávila Ávila

Carlos Alberto Quintero Sánchez

Prof. Carlos Alberto Franco Franco

Administración en Logística y Producción

Administración de Empresas

Bogotá D.C

2023

### **Declaración de originalidad y autonomía**

Declaro bajo la gravedad del juramento, que he escrito el documento de título “Proyecto de Investigación Capstone “Tucan S.A.””, en la opción de grado de Capstone y que por lo tanto, su contenido es original.

Declaro que he indicado clara y precisamente todas las fuentes directas e indirectas de información y que este trabajo no ha sido entregado a ninguna otra institución con fines de calificación o publicación.

Isabella Ávila Ávila

Carlos Alberto Quintero

Daniela Tole Silva

**Declaración de exoneración de responsabilidad**

Declaro que la responsabilidad intelectual del presente trabajo es exclusivamente de su autor. La Universidad del Rosario no se hace responsable de contenidos, opiniones o ideologías expresadas total o parcialmente en él.

Isabella Ávila Ávila

Carlos Alberto Quintero

Daniela Tole Silva

## Tabla de contenido

Resumen.....	8
Palabras clave.....	11
Abstract.....	12
Key Words .....	12
<b>1. Introducción.....</b>	<b>13</b>
<b>2. Objetivos.....</b>	<b>15</b>
<b>3. Acerca de Tucan .....</b>	<b>15</b>
<b>4. Metodología.....</b>	<b>16</b>
<b>4.1. Trabajo de campo.....</b>	<b>16</b>
<b>4.2. Datos e información otorgada Tucan Flowers S.A.....</b>	<b>22</b>
<b>4.2.1. Variedades de Flor.....</b>	<b>22</b>
<b>4.2.2. Enfermedades de las plantas y producción.....</b>	<b>23</b>
<b>4.2.3. Promedio de horas trabajadas. (Mes de marzo) .....</b>	<b>24</b>
<b>5. Análisis de producción de mini rosas para Tucan Flowers S.A. ....</b>	<b>28</b>
<b>5.1. ¿Qué es un análisis de datos? ¿Cuál es su utilidad? .....</b>	<b>28</b>
<b>5.2. ¿Por qué Python? .....</b>	<b>28</b>
<b>5.3. Desarrollo de los códigos en Python .....</b>	<b>29</b>
<b>5.4. Desarrollo del análisis final en Excel.....</b>	<b>33</b>
<b>6. Resultados.....</b>	<b>34</b>
<b>6.1. Resultados de la producción de mini rosas de Tucan 3 para el año 2022 .....</b>	<b>34</b>
<b>6.1.2. Análisis general 2022 .....</b>	<b>35</b>
<b>6.1.3. Análisis descriptivo de cada bloque 2022 .....</b>	<b>36</b>
<b>6.1.4. Análisis específico 2022.....</b>	<b>37</b>
<b>6.2. Resultados de la producción de mini rosas de Tucan 3 para el año 2023 .....</b>	<b>38</b>
<b>6.2.1. Análisis general 2023.....</b>	<b>39</b>
<b>6.2.2. Análisis descriptivo de cada bloque.....</b>	<b>40</b>
<b>6.2.3. Análisis específico .....</b>	<b>41</b>
<b>6.3. Reunión concluyente con el agrónomo .....</b>	<b>42</b>
<b>6.3.1. Recomendaciones Empresariales .....</b>	<b>43</b>
<b>7. Conclusiones.....</b>	<b>44</b>
<b>8. Bibliografía.....</b>	<b>46</b>

**Indice de imágenes**

Imagen 1 Ubicación de cultivos y oficinas de Tucan .....	17
Imagen 2 División de hierbas y sistemas de poleas Tucan Flowers S.A. ....	18
Imagen 3 Rosas en las cajas y respectivas estampas de clasificación .....	19
Imagen 4 Poscosecha: Selección, corte y armado .....	20
Imagen 5 Ramos armados, cajas dependiendo del exportador y rosas provenientes de otras fincas .....	21

## Índice de ilustraciones

Gráfico 1 Desviación Estándar de la Producción del Bloque 1 durante el año 2022 .....	38
Gráfico 2 Diagrama de cajas y bigotes para el Bloque 1 de Tucan 3 durante el año 2022 .....	38
Gráfico 3 Organización del análisis específico del bloque 1 de la producción de Tucan 3 para el año 2022.....	39
Gráfico 4 Diagrama de cajas y bigotes para el Bloque 1 de Tucan 3 durante el año 2023 .....	41
Gráfico 5 Gráfica de organización del análisis del bloque 1 de la producción de Tucan 3 para el 2023.....	42

**Indice de Tablas**

Tabla 1 Captura documento entregado por Tucan .....	25
Tabla 2 Captura del Archivo “FORMATO PROMEDIO X HORAS TRABAJADAS MARZO 01 DE 2023.....	26
Tabla 3 Datos Tucan Finca #3 .....	28
Tabla 4 Datos Organización de Producción 2022.....	28
Tabla 5 Tabla de estadística descriptiva del bloque 1.....	37
Tabla 6 Tabla de estadística descriptiva del bloque 1 2023.....	41

## Glosario

**Variedad:** Diferentes tipos de rosas con características específicas.

**Injerto:** Técnica de unir dos plantas para combinar sus características.

**Portainjertos:** Plantas utilizadas como base para el injerto de rosas.

**Invernadero:** Estructura cerrada para controlar el ambiente de cultivo.

**Vivero:** Área de preparación y crecimiento de las plantas jóvenes.

**Sistema de Riego:** Métodos para proporcionar agua de manera controlada.

**Poda:** Recorte de ramas para favorecer el crecimiento y la floración.

**Fertilización:** Aporte de nutrientes esenciales para el desarrollo de las rosas.

**Control de Plagas:** Medidas para prevenir y tratar problemas con insectos u enfermedades.

**Cosecha:** Recolección de rosas en el momento adecuado.

**Clasificación:** Separación de las rosas según su tamaño y calidad.

**Empaque:** Preparación de las rosas para su transporte y venta.

**Transporte:** Movimiento de las rosas desde la finca hasta los puntos de venta.

**Almacenamiento:** Conservación de las rosas antes de su distribución.

**Cadena de Suministro:** Sistema que abarca la producción, distribución y venta.

**Automatización:** Uso de maquinaria y sistemas para optimizar procesos.

**Monitoreo:** Seguimiento constante de condiciones climáticas y de cultivo.

**Sistemas de Información:** Tecnologías para gestionar datos sobre la producción

**Control de Calidad:** Procedimientos para asegurar que las rosas cumplen con estándares establecidos.

**Python:** Un lenguaje de programación versátil y de alto nivel utilizado ampliamente en ciencia de datos y análisis

**DataFrame:** Estructura de datos tabular de Pandas, similar a una hoja de cálculo.

**Estadísticas Descriptivas:** Medidas que resumen y describen características del conjunto de datos.

**Histograma:** Representación gráfica de la distribución de un conjunto de datos.

**Diagrama de Caja (Boxplot):** Gráfico que muestra la distribución de un conjunto de datos y sus estadísticas resumidas.

**Pruebas Estadísticas:** Métodos para realizar inferencias sobre poblaciones basadas en muestras.

**Correlación:** Medida de la relación entre dos variables.

**Análisis de Tendencias a lo Largo del Tiempo:** Utilización de gráficos y tablas para analizar cómo ciertos aspectos del cultivo han cambiado a lo largo de varias temporadas.

**Evaluación de Rendimiento por Variedad:** Hoja de cálculo que compara el rendimiento de diferentes variedades de rosas en términos de crecimiento, floración y resistencia.

## **Resumen**

El sector agrícola es uno de los más retadores en cuanto a procesos de producción y logística debido a que se enfrenta a múltiples variables que no son fáciles de controlar como lo son el clima, el manejo de plagas y enfermedades, el estado de las vías, entre otros. En países poco desarrollados donde no se cuenta con la tecnología suficiente para estandarizar y automatizar procesos, como en el caso de Colombia, se adiciona la variable de mano de obra la cual es poco capacitada y escasa. Por tal motivo, es difícil generar pronósticos de cosechas confiables que se ajusten con las proyecciones de venta. Una de las soluciones consiste en la trazabilidad de los procesos que permita identificar factores claves para poder corregirlos y cumplir con los objetivos de producción y venta.

Tucan Flowers es una empresa productora de mini rosas que es reconocida por sus altos estándares de calidad en mercados internacionales como Estados Unidos, pero que se enfrenta a los problemas mencionados anteriormente. Por ende, en el siguiente trabajo se realizó un diagnóstico de su producción de mini rosas desde el año 2022 hasta agosto del 2023, para identificar factores claves que afectan sus procesos productivos. Estos fueron: la estacionalidad y la productividad de las variedades de mini rosas. Se usaron herramientas como Python y Excel. De igual forma, sienta las bases para el desarrollo de futuros proyectos que diseñen modelos de producción con los ajustes en cada uno de estos factores claves ya identificados para, predecir la cosecha y cumplir con los objetivos de venta.

### **Palabras clave**

Análisis multivariable, pronóstico, procesos, floricultura, factores clave

## **Abstract**

The agricultural sector is one of the most challenging in terms of production processes and logistics due to facing multiple variables that are not easy to control, such as climate, pest and disease management, and the condition of roads, among others. In underdeveloped countries where there is insufficient technology to standardize and automate processes, as in the case of Colombia, the variable of labor is added, which is both unskilled and scarce. Therefore, generating reliable harvest forecasts that align with sales projections is difficult. One solution involves traceability of processes to identify key factors for correction and to meet production and sales objectives.

Tucan Flowers is a mini rose-producing company known for its high-quality standards in international markets such as the United States, but it faces the problems. Consequently, this study conducted a diagnosis of its mini rose production from 2022 to August 2023 to identify key factors affecting its production processes. These factors included seasonality and the productivity of mini rose varieties. Tools such as Python and Excel were used. Similarly, it lays the groundwork for the development of future projects that design production models with adjustments in each of these identified key factors to predict harvests and meet sales objectives.

## **Key Words**

Multivariate analysis, forecasting, processes, floriculture, key factors.

## 1. Introducción

La industria de las flores es una de las más influyentes actualmente en el PIB nacional, específicamente industria agrícola en el país, representando el segundo renglón de exportaciones agrícolas en términos de valor (Ministerio de Agricultura, 2021). En términos específicos de exportación, Colombia es el segundo exportador más grande de flores a nivel mundial, siendo el principal cliente Estados Unidos, seguido de Canadá, Reino Unido, Holanda y Jamaica. Si bien se considera que el mercado nacional es sumamente fructífero, según el Ministerio de Agricultura y Desarrollo Rural este sector en particular tiene un enfoque de desarrollo en el mercado internacional, siendo los principales medios de transporte el aéreo y el marítimo.

En cuanto a la cosecha de flores, los departamentos líderes en la industria son Cundinamarca con un 66% en la producción nacional, seguido de Antioquia con un 33% y por último departamentos de la zona centro occidente como Risaralda, Caldas, Quindío y Valle del Cauca representan un 1% (Ministerio de Agricultura, 2021). Dentro de las especies que más se exportan se encuentran la rosa, el clavel, el crisantemos, la hortensia y la alstroemerias, y las temporadas del año en las que se realizan más exportaciones son San Valentín, el día de la madre y por último, el día de acción de gracias.

En términos del sector agrícola, es fundamental mencionar que se encuentran 3,14 millones de trabajadores desempeñando labores desde riego, cosecha hasta distribución local de diferentes productos, siendo esto un 17% del total de la población colombiana. (Statista, 2023). Este panorama no sólo evidencia un pilar económico fundamental para la economía colombiana de carácter interno y externo, sino que también social puesto que las familias que se sostiene bajo esta industria y los alimentos de primera necesidad son el resultado de la producción de procesos agrícolas, y que hoy gracias al crecimiento desacelerado de la

población tienen más demanda que antes, priorizando el factor de la seguridad alimenticia para todos. (T. Rehman, C. Romero;1993).

Al tener conocimiento de las variedades que más se exporta y en las temporadas en donde más se exportan, se creería que es sencillo realizar un diagnóstico de la demanda, sin embargo, muchas compañías dentro de la industria de las exportaciones y multinacionales no cuentan con los datos y las herramientas suficientes para realizar un registro y un posterior pronóstico de la demanda, o en un escenario más crítico como expone Konstantinos Perakis en su texto “CYBELE” en el que puede que las compañías tengan las herramientas, pero se desconoce la funcionalidad, no se tiene conocimiento para la interpretación de datos para distintos niveles organizacionales de la compañía (Perakis et al., 2020) , dando como resultado no sólo pérdidas de carácter económico, sino también social tal y como se aclaró previamente. Un escenario de tal índole es el de la compañía Tucan S.A., empresa colombiana de producción y exportación de flores, la cual durante estas fechas recibe grandes pedidos, principalmente para exportar a Estados Unidos y que ha manifestado inconvenientes a la hora de realizar el proceso de cultivo, mantenimiento y envío de mini rosas.

Con el fin de prever el impacto en los procesos de la compañía, Tucan S.A. decidió trabajar de la mano con la Universidad del Rosario para desarrollar un proyecto que tenga como finalidad elaborar un análisis de la demanda de la compañía en cuestión, tomando los datos de registro y la producción de una de las tres fincas que posee la compañía, específicamente la finca #3. De manera puntual, el propósito de este documento es enunciar el trabajo realizado durante el primer semestre del 2023 por parte del equipo de analítica de datos de la empresa Tucan S.A., liderado por el profesor Carlos Alberto Franco Franco y conformado por los estudiantes de la Escuela de Administración Daniela Tole, Isabella Avila y Carlos Alberto

Quintero, el cual tiene como principal objetivo predecir la demanda de mini rosas para exportación a países como Estados Unidos y España.

## **2. Objetivos**

Priorizando el factor de trabajo en conjunto integral, junto con la compañía Tucan S.A. se abrió un espacio para compartir las opiniones y expectativas que se tenía sobre el proyecto, siendo el resultado en conjunto los siguientes objetivos:

- Conocer y relacionarse con la base de datos utilizada por la compañía Tucan S.A., específicamente de la base utilizada en la finca #3 de la compañía durante los años 2022 y el año 2023, este último hasta el mes de agosto.
- Utilizar herramientas de análisis de datos, específicamente el programa Python para sintetizar, observar y analizar el comportamiento de la producción según determinadas variables.
- Interpretar los datos obtenidos y generar recomendaciones para la contribución a la potencial mejora de procesos productivos de Tucan S.A.

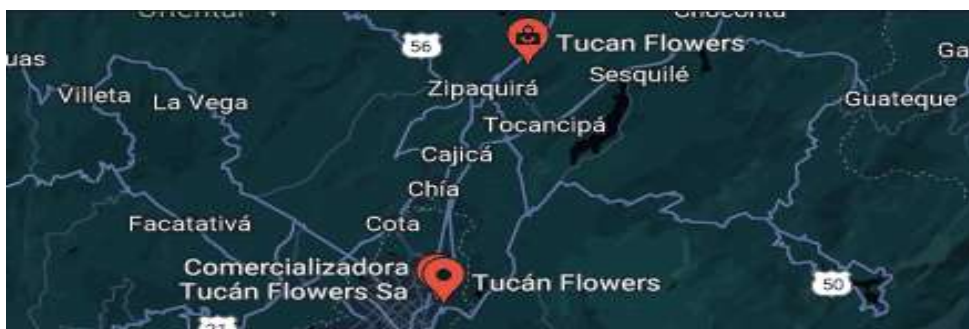
## **3. Acerca de Tucan Flowers S.A.**

Tucan Flowers S.A. es una compañía que se dedica a realizar el proceso de cosecha, distribución y exportación de rosas y mini rosas. La compañía fue fundada hace más de 18 años y desde entonces ha sido liderada por la familia Contreras, siendo el gerente actualmente Gustavo Contreras. Hoy en día, la compañía tiene 200 empleados y las 3 fincas que se tienen

destinadas para los cultivos están ubicadas al norte de Bogotá, específicamente entre los municipios de Cogua y Nemocón. Actualmente el mayor mercado al que Tucan exporta es el estadounidense, es más, se considera que aproximadamente de toda la producción, el 80% es exportado.

### **Imagen 1**

*Ubicación de cultivos y oficinas de Tucan Flowers S.A.*



Fuente: Google Maps, 2023.

## **4. Metodología**

### **4.1. Trabajo de campo**

En un principio se discutió la necesidad de un primer acercamiento formal entre el equipo de la Universidad del Rosario y la compañía Tucan, adicional al ya hecho en las entrevistas para hacer parte del Proyecto Capstone 2023. Para esa salida se realizó un agendamiento para la visita de una de las fincas ubicadas entre los municipios de Cogua y Nemocón, en la que se realizó una visita a los diferentes cultivos y nos presentaron todos los

procesos que se realizan dentro de cada uno de ellos. El día 10 de marzo de 2023, se realizó la visita a la finca denominada Tucan 1, la cual se considera es la finca principal en la que se desarrolla el procesamiento, cosecha y envíos de las plantas, mientras que en las fincas Tucan 2 y Tucan 3 se realiza exclusivamente la cosecha y el producto se transporta a la finca Tucan 1 para el proceso de postcosecha y debida preparación.

En Tucan 1, como se mencionó previamente se realiza un proceso de cosecha, desarrollado en diferentes invernaderos y dividido por tipo de rosas.

## **Imagen 2**

*División de hileras y sistema de poleas Tucan Flowers S.A.*



Fuente: elaboración propia.

En términos de proceso, la cosecha se hace acorde a la necesidad de la demanda, la cual puede ser germinación por semilla o por injerto y son tratados y divididos en términos de trabajo por naves y camas, siendo estructuras elevadas metálicas e hileras de cosecha. Una vez pasa el tiempo de cosecha, el cual es aproximadamente de 2 a 3 meses, son cortadas por trabajadores y colocadas en cajas, las cuales tienen un registro y clasificación acorde a la finca, invernadero, el bloque, la cama, el tallo, tipo/color, semana y fecha.

**Imagen 3**

*Rosas en las cajas y respectivas estampas de clasificación*



Fuente: elaboración propia.

Posteriormente son desplazadas por un sistema de poleas a través de la finca hasta la zona de postcosecha, en la que un grupo de trabajadores revisan la calidad de la rosa, cortan los tallos y preparan los ramos acordes a las órdenes desde dirección general y es en este punto en donde convergen todas las rosas provenientes de las múltiples fincas.

**Imagen 4**

*Postcosecha: Selección, corte y armado.*



Fuente: elaboración propia.

Posteriormente son registrados y almacenados en un cuarto frío, el cual está clasificado acorde al pedido y fecha de salida. De igual forma, en este espacio se encuentran las cajas están etiquetadas acorde las especificaciones de los clientes tanto a nivel nacional como internacional, para luego ser distribuidas en camiones hasta los centros de distribución y posterior exportación en el Aeropuerto el Dorado, ubicado en Bogotá D.C.

### **Imagen 5**

*Ramos armados, cajas dependiendo del exportador y rosas provenientes de otras fincas.*



Fuente: elaboración propia.

A continuación, se encuentran los datos de un mapa de procesos que sintetiza lo observado en la visita previamente descrita. Se realizó el diagrama con la finalidad de tener claros los procesos y de cómo potencialmente los datos de las diferentes etapas influyen en la proyección final

**Figura 1**

*Mapa del proceso de Tucan Flowers S.A.*



## **4.2. Datos e información otorgada Tucan Flowers S.A.**

### ***4.2.1. Variedades de Flor***

Tucan cuenta con un total de cuarenta y cinco (45) variedades, cada una cuenta con un código, el color, los grados y el tipo de género de flor. Con estos datos se busca analizar cuál es la variedad que más se cosecha, cuáles son las que presentan mayor número de problemas (enfermedades) y lograr predecir en qué época del año es oportuno cosechar cada tipo de flor.

Es preciso tener en cuenta que este proceso tiene un control riguroso en términos de calidad que va desde control de plagas hasta conservación y selección de las rosas acorde a la demanda, y es precisamente en este punto en donde el equipo empieza su trabajo. El propósito de este proyecto es por medio de un análisis de datos, programado por medio de la plataforma Python realizar recomendaciones con respecto a qué procesos deben llevarse a cabo o que tipo de configuración se debe presentar dentro del ya existente, acorde a la potencial demanda que se pueda encontrar con los datos con los que en la actualidad la compañía trabaja para su producción.

Posterior a la visita inicial y durante todo el desarrollo del proyecto, el profesor Carlos Alberto Franco Franco fue el puente directo de comunicación entre el equipo de trabajo y Gustavo Contreras, tal y como fue acordado para una mayor fluidez durante el proceso de trabajo. Al principio, el profesor Carlos Franco solicitó documentos de registro de años recientes sobre factores que afectan la producción de una finca en particular, tomando con fines de eficiencia la finca número 3 que era la que poseía mayores datos y posteriormente se planteó la siguiente agenda de trabajo:

Los datos que fueron enviados inicialmente por el área de compras y ventas fueron los siguientes:

- Categorización del proceso: En este inciso el propósito inicial era identificar los datos, realizar una organización que permitiera tanto al equipo, al programa Python como a las personas de la compañía comprender mejor los datos. La fecha estimada de realización fue establecida para ser hecha entre los meses de abril y mayo de 2023.

- Predicción: Direccionado a cómo cada variable del análisis estadístico planeado inicialmente puede influir en el proceso de producción. Fecha estimada sería los meses de agosto y septiembre de 2023.

Una vez planteado el esquema de trabajo, se clasificaron los datos enviados inicialmente por el área de compras en:

#### ***4.2.2. Enfermedades de las plantas y producción***

Este grupo de documentos comprende información sobre la cantidad de cortes realizados en los meses de enero y febrero y las enfermedades presentadas en cada variedad para los mismos meses, además de la producción total por variedad para el año 2022. Con esta información, se busca analizar cuáles son las variedades con mayor producción y observar la relación que puede existir entre las enfermedades que afectan a estas variedades con la producción final.

En un comienzo, se subió la información a Python, pero el programa no pudo leer los datos debido a que no están estandarizados, es decir, no todas las hojas del excel estaban organizadas en el mismo formato. Por lo que no se pudo trabajar con ellos. Como

recomendación, se sugiere que los datos tengan un formato estandarizado para poder utilizarlos en programas de data analítica.

- Promedio de horas trabajadas. (Mes de marzo)
- Variedades de flores.

**Tabla 1**

*Captura documento de excel Tucan 3 2022*

Variedad	Abierto	Torcido	Velloso	Polvoso	Botritis	Araña	Maltrato	corto	2cabez/ des	Afidos
WHITE MAYOLIKA	86	8	0	56	0	27	11	77	10	0
LOVELY LIDIA	31	6	0	38	0	29	10	34	8	0
SUNCITY	19	2	0	17	0	5	3	10	7	0
ILSE	87	9	0	56	0	49	10	82	8	0
SUPERNOVA	55	6	0	50	0	43	10	26	7	0
SCARLETTE MIMMI	34	5	0	29	0	28	9	26	6	0
SENTHYNA	35	2	0	13	0	5	3	12	2	0
MAMBO	28	3	0	25	0	12	3	5	2	0
RUMBA	19	2	0	13	0	10	3	4	2	0
PINK MAYOLIKA	90	8	0	51	0	36	7	79	5	0
BABE	54	7	0	39	0	21	8	16	7	0
VIVI	0	0	0	0	0	0	0	0	0	0
SNOW FLAKE	67	5	0	40	0	27	7	17	4	0
PORCELINA	0	0	0	0	0	0	0	0	0	0
LIDIA	51	4	0	41	0	35	5	7	4	0
ZABRA	32	3	0	22	0	12	5	8	3	0
CRETA	0	0	0	0	0	0	0	0	0	0
CELESTIAL	0	0	0	0	0	0	0	0	0	0

Fuente: Tucan Flowers (2023). *documento de excel Tucan 3 2022*.

#### **4.2.3. Promedio de horas trabajadas. (Mes de marzo)**

El área de compras y ventas nos envió el registro del mes de marzo de la productividad por trabajador en la franja horaria de 6 a.m. a 3 p.m, teniendo como variable central cantidad de rosas preparadas. El formato estándar que se manejó para el mes de marzo fue el siguiente:

**Tabla 2**

Captura archivo "FORMATO PROMEDIO X HORAS TRABAJADAS MARZO DE 2023

	6 - 7 A	7 - 8 A	8 - 9 A	9 - 10 A	10-11 a	11-12 p	12-01 P	01-02 P	02-03 P	UBTOT	PROMEDI
<b>Yeimi Diaz/Nicol Mondragon</b>	180	214	206	409	434	169	368	398	180	2618	308
<b>Yanet Lozano</b>		168	40	215	234	90	30	100	156	1033	122
<b>Yorleidy Perez/Paola Pulido</b>	310	372	280	454	360	199	440	240	586	3241	381
<b>Keidy Pirela</b>	170	190	211	252	158	47	225	63	146	1462	172
<b>Ximena Nieto</b>		248	166	189	162		180	109	135	1189	140
<b>Nelly Barrera/Adriana Acosta</b>	220	280	270	342	291	164	270	243	486	2566	311
<b>Luz Mery Bachiller/Adriana Zan</b>	180	370	334	437	330	170	236	200	468	2785	338
<b>Marta Castro/Yuranis Ibarra</b>	120	180	320	290	464	120	380	220	762	2856	336
<b>Karla Espinosa</b>	40	100	90	170	170	70	118	70	170	998	117
<b>Adelis Alayon/Norexi Belmonte</b>		281	244	282	235		361	270	180	1853	225
<b>Yesika Oquendo</b>	60	200	120	180	95	40	166	156	118	1135	134
<b>Carmen Miranda/Karen Jades</b>	230	300	320	400	300	200	392	389	383	2914	353
<b>Miriam Montero/Dayana Lopez</b>	120	322	130	330	218	110	274	237	217	1958	230
										0	0

Fuente: Tucan Flowers. (2023). *FORMATO PROMEDIO X HORAS TRABAJADAS MARZO DE 2023*.

En este formato se puede entonces identificar de forma sencilla por medio del subtotal y del promedio la productividad de los trabajadores, y en un análisis preliminar se pudieron identificar tres cosas. La primera es que las horas de mayor productividad son entre las 9 a.m. y las 11:00 a.m. y segundo, que la productividad entre individuos/parejas de trabajo es variable en el transcurso del mes. Sin embargo, se puede destacar que Carmen Miranda/Karen Jades y Adelis Layon/Norexi Belmo son de las duplas que presentan mayores números en términos de tratamiento de las rosas. Es preciso denotar que estas observaciones son de carácter preliminar, por lo que se pueden terminar cambiando en cierto grado y luego tratamiento de datos con el programa Python, los análisis presentados pueden discrepar. A pesar de realizar un análisis general de los resultados, el equipo se percató que no era posible realmente realizar un ejercicio de programación que permitiera generar una relación

En esta parte del trabajo culminó el trabajo del semestre 2023-01, pero conversando en espacios extracurriculares el equipo se percató que, con los datos recolectados y la disposición realizada de los mismo, no era posible un manejo de datos para realizar un proceso de programación y análisis estadístico como se tenía planteado desde un inicio porque se encontraron múltiples vacíos de información y era complejo cubrir toda la información para realizar un trabajo íntegro. Por este motivo, se preguntó a la compañía si era posible realizar una visita a una de las fincas en búsqueda de otro tipo de datos con un enfoque específico en la producción, a lo que nos asignaron una visita para el día 9 de agosto.

La visita realizada el 9 de agosto se discutió por qué los datos previamente enviados no eran los más indicados para llevar a cabo el proyecto, por lo que con dos agrónomos encargados de los procesos internos de Tucan S.A. y nos sugirieron trabajar con documentos de registro de la producción, específicamente de la finca #3 de los años 2022 y 2023, ya que fueron años donde se realizó el registro de una forma más ordenada. Inicialmente, cada hoja del archivo representaba un mes, el cual contenía las variedades de mini rosas por bloque y su producción diaria, semanal y total.

Sin embargo, esta organización no era útil para Python, dado que este programa no lee celdas unidas ni en blanco y siempre empieza a leer desde la primera celda de cada columna y fila. Por lo tanto, se decidió agrupar la información por bloques de la siguiente forma:

### **Tabla 3**

*Producción Tucan 3 2023*



## **5. Análisis de producción de mini rosas para Tucan Flowers S.A.**

### **5.1. ¿Qué es un análisis de datos? ¿Cuál es su utilidad?**

"Actualmente, existen tantas fuentes de información que damos por sentado el valor de los datos que la gente otorga cuando interactúa con un anuncio, cuando realiza una compra o cuando comparte una opinión sobre un producto" (Pursell, s. f.).

Día a día, las compañías intercambian, generan y recolectan miles de datos, no obstante, hoy en día no solo basta con realizar estas acciones, sino que es indispensable el saber analizarlos para poder entender las dinámicas del mundo y los procesos internos en las compañías y, de esta forma, ofrecer un servicio que se ajuste a las necesidades del cliente. En esto consiste la analítica de datos, en saber interpretar y organizar los datos para que se puedan tomar las decisiones correctas.

### **5.2. ¿Por qué Python?**

Se consideró el uso de la herramienta de programación Python porque es una herramienta útil para trabajar con datos grandes y extensos presentes en distintos formatos. De

igual forma, permite hacer análisis más robustos y de forma mucho más rápida, dentro de los que se destaca la limpieza de datos, es decir, reemplazarlos por un valor determinado. Python igualmente posee una extensa variedad de opciones para visualización de datos como gráficos y tablas. Finalmente ofrece librerías para análisis de datos, brinda otras herramientas como la creación de bases de datos.

### 5.3. Desarrollo de los códigos en Python

Luego de que los datos se organizaron en Excel, se exportaron a Spyder, un programa de Python, en el cual se construyeron los códigos necesarios para poder llevar a cabo dos tipos de análisis: descriptivo y multivariable. Antes de explicar la metodología empleada en este paso, es importante tener en cuenta los siguientes conceptos básicos:

- Librerías: “ conjuntos de funciones que permiten realizar distintas funciones” (Chaves, 2022).

- Pandas: “ La biblioteca de software de código abierto Pandas está diseñada específicamente para la manipulación y el análisis de datos en el lenguaje Python” (alexandre, 2022).

- Dataframe: “son paneles bidimensionales compuestos por filas y columnas, que permiten destacar las relaciones entre las distintas variables de la serie de datos.” (Team, 2022).

- Tipos de datos usado en Python

Python puede usar una gran cantidad de tipos de datos, no obstante, para este proyecto solo se usaron dos:



El procedimiento en Spyder se puede dividir en tres partes: Primero, la lectura de los datos y el uso de librerías como Pandas y matplotlib para el desarrollo del análisis descriptivo. Segundo, el uso de librerías como numpy para el análisis multivariable; y tercero, la descarga de los datos en archivos XX y de Excel. Vale la pena tener en cuenta que este procedimiento es el mismo tanto para los datos del año 2022 como para los del 2023.

Para el primer paso, se tuvieron en cuenta funciones como Print(), head(), info(), insull().sum(), describe(include= 'all') y plt.boxplot(). No obstante, la más importante fue describe(include= 'all'), la cual permite obtener las siguientes medidas estadísticas que se usaron en análisis posteriores:

- Count: dice cuántos datos hay por columna
- Unique, Top y Freq son medidas para los datos de denominación string (palabras). Unique es la cantidad de valores que no se repiten, mientras que Top es el valor más común y Freq es su frecuencia.
- Mean: la media.
- Std: desviación estándar
- Min: El valor mínimo
- 25%: Los datos que están por igual o por debajo del 25%.
- 50%: los datos que están por igual o por debajo del 50%.
- 75%: los datos que están por igual o por debajo del 70%.
- Max: el valor máximo.

En la segunda parte, el objetivo principal era realizar un análisis de la producción de acuerdo con el mes, semana y día. Para ello se hizo uso de las siguientes dos funciones:

```
print(.groupby(['MES','VARIEDAD'])['PRODUCCIÓN'].sum() )
```

```
print(.groupby(['MES','VARIEDAD'])['PRODUCCIÓN'].mean() )
```

De este modo, cada función organizaba las tres variables mencionadas “MES”, “VARIEDAD” Y “PRODUCCIÓN” en un solo dataframe, no obstante, la diferencia de una función con la otra radica en que la primera organiza los datos totales, mientras que la segunda los promedia. Partiendo de estas tablas se puede saber cuántas mini rosas de la variedad x, se produjeron en el mes y.

Este procedimiento se repite para cada bloque y para los análisis realizados por semana y día.

En el tercer paso, se realizaron códigos más extensos para descargar los datos en dos formatos: el primer tipo texto y el segundo xlsx, es decir, archivo de Excel.

Para la primera descarga, se convirtieron los datos con las siguientes tres funciones:

```
dataframe= open('nombre', "a")  
dataframe.write(str(función))  
dataframe.close()
```

Este procedimiento se hizo por cada bloque. La segunda función se empleó para cada función obtenida en cada uno de los análisis.

Posteriormente, se decidió realizar un solo dataframe de texto. Por lo cual, se unieron los resultados anteriores de cada bloque con la siguiente función:

```
dataframe = pd.merge(left= ,right= , how = 'outer')
```

La función pd.merge solo permite unir dos tablas, por lo que para se tuvo que ir uniendo de a dos bloques hasta llegar a una sola tabla.

Para la segunda descarga, que fue a un archivo en Excel, se realizó el siguiente procedimiento:

Se creó una dataframe para cada código de cada bloque con la siguiente función: `dataframe = pd.DataFrame()`. Luego, cada dataframe creada se utilizó en el siguiente código:

```
with pd.ExcelWriter('.xlsx', engine='xlsxwriter') as writer:  
    dataframe.to_excel(writer, sheet_name='name').
```

#### **5.4. Desarrollo del análisis final en Excel**

El último paso para la construcción de los análisis fue la organización de los datos obtenidos en Python en tablas dinámicas. Cada año se desarrolló en un archivo diferente, no obstante, siguen la misma estructura como se muestra a continuación:

##### **Figura 3**

*Análisis de la producción de Tucan 3 para el año 2022 en archivo excel*



Fuente: elaboración propia

La primera hoja consiste en un análisis general del total de la producción de Tucan 3, mientras que las siguientes hojas corresponden a cada uno de los análisis descriptivos y multivariantes de los bloques.

## 6. Resultados

### 6.1. Resultados de la producción de mini rosas de Tucan 3 para el año 2022

Los resultados se pueden dividir de acuerdo con los análisis realizados y los bloques. En el anexo XX se realizó un análisis general, que abarca la producción de todas las variedades de mini rosas y la producción total de la finca; y uno específico por cada bloque en el cual se observa detalladamente la producción por mes y semana, cada uno en términos de producción total y promedio. En el presente documento, se describe cómo están organizados estos

resultados para el Bloque 1 y el Análisis general, dado que los análisis de los demás bloques ya se encuentran en el anexo correspondiente.

Sin embargo, es importante resaltar que, para el año 2022 se omitió el análisis del Bloque 9 dado que no hubo cosechas en ningún momento del año. Además, se omitió el análisis de la producción diaria, dado que, para los primeros seis meses del año, no se encontraba producción diaria, por el contrario, una gran cosecha en un solo día, que solía ser el último de cada mes. Por ende, sería arbitrario tener en cuenta esta variable, ya que el resultado no sería acertado.

### ***6.1.2. Análisis general 2022***

En este, se crearon tres tablas con sus respectivas gráficas, cada una con un objetivo diferente. Estas son:

- Producción total de Tucan 3 durante el año 2022: este gráfico recopila la producción total de cada mes independientemente del mes. Permite observar patrones de cosecha a nivel general.
- Total y Promedio de Producción por bloque: en él, se observa la productividad por bloque. Su utilidad radica en que permite observar cuánto produce cada bloque al año y qué tan uniformes son sus datos. Es decir, si su promedio es bajo y la producción total alta, significa que hay pocas cosechas grandes pero que en el día a día se cosechan pocas mini rosas.
- Producción total por variedad: permite observar cuáles son las variedades de mini rosas que más producen independientemente del bloque en el que se encuentren.

### 6.1.3. Análisis descriptivo de cada bloque 2022

Este fue el primer análisis trabajado en Python y reúne una tabla con las principales medidas estadísticas como se muestra a continuación:

**Tabla 5**

*Estadística descriptiva del bloque 1*

	BLOQUE	DÍA	SEMANA	MES	VARIEDAD	PRODUCCIÓN
Número de filas	1092	1.092	1.092	1092	1092	1.092
Media	1	16	26	0	0	784
Desviación Estándar	0	9	15	0	0	3.079
Valor Mínimo	1	1	1	0	0	0,00
25%	1	8	14	0	0	0,00
50%	1	16	27	0	0	0,00
75%	1	23	39	0	0	861
Valor Máximo	1	31	52	0	0	41.808

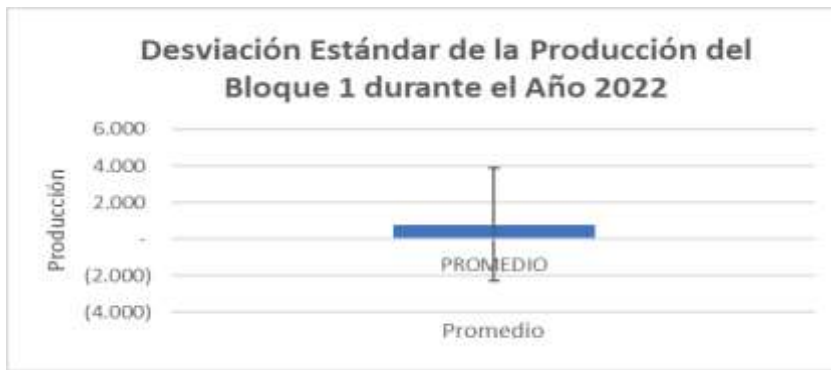
Fuente: elaboración propia

Como se observa, para las tablas de cada bloque se excluyeron las medidas de Unique, Top and Freq que se obtenían con uno de los códigos de Python, dado que estas hacían referencia a variables no numéricas que resultaban poco útiles para los análisis.

Luego, se realizó una gráfica de desviación estándar que permitía observar qué tan variables eran los datos.

### **Gráfico 1.**

*Desviación estándar de la producción del Bloque 1 durante el año 2022*

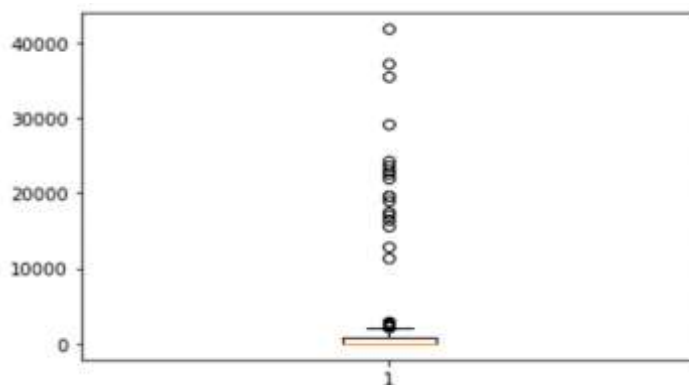


Fuente: elaboración propia

Por último, se incluyen los gráficos de Cajas y Bigotes que cada código generaba. Es importante tener en cuenta que no todos los bloques del 2022 tienen un gráfico ya que sus datos son muy pocos o dispersos, por lo cual no podía generarlos.

## Gráfico 2.

*Diagrama de cajas y bigotes para el Bloque 1 de Tucan 3 durante el año 2022*



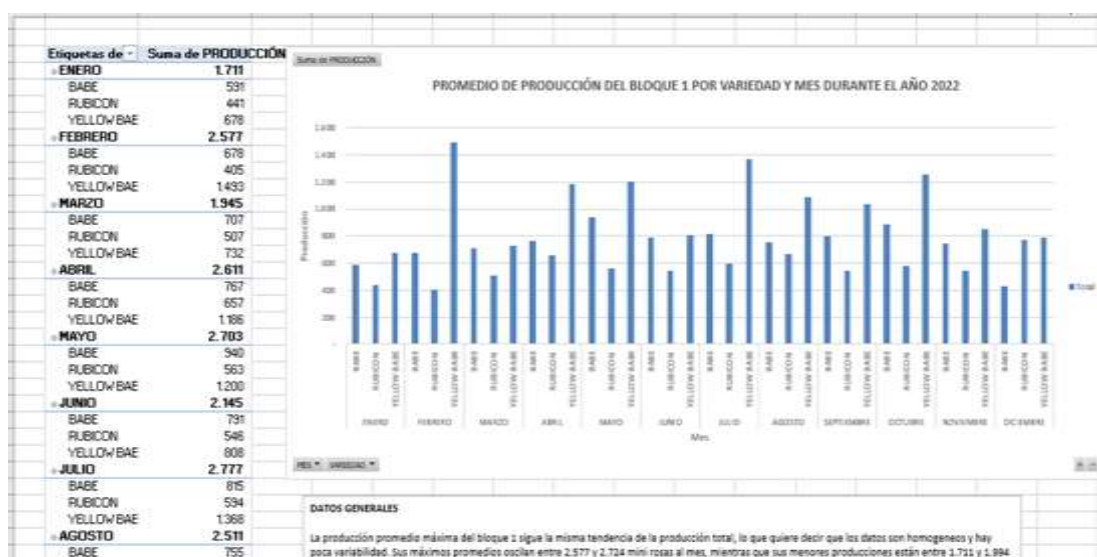
Fuente: elaboración propia

### 6.1.4. Análisis específico 2022

Cada grupo de datos ya sea por mes o por semana, está organizado en una tabla dinámica que permite ver a detalle no solo el comportamiento de todo el bloque, sino por variedad. Cada hoja tiene su propio análisis en el cual se resaltan los insights más importantes a nivel general y las variedades más relevantes.

### Gráfico 3

Organización del análisis del bloque 1 de la producción de Tucan 3 para el 2022



Fuente: elaboración propia

## 6.2. Resultados de la producción de mini rosas de Tucan 3 para el año 2023

Así como se desarrolló en el inciso del año 2022, para el año 2023 fueron divididos acorde a los análisis realizados y los bloques. En primer lugar, se realizó un análisis general, el cual abarca la producción de todas las variedades de mini rosas y la producción total de la finca,

y un análisis más específico por cada bloque presenta (del 1 al 9), en el que se observa de forma detallada la producción por mes, semana y día, cada uno en términos de producción total y en promedio. En el presente documento se describe cómo están organizados estos resultados para el bloque 1 y el Análisis General, ya que por cuestiones de formalidad y orden el resto de los procedimientos se encuentra en el anexo 3. Es pertinente mencionar que se tuvo en cuenta el rango de tiempo desde enero hasta agosto, dado que se tenía registro hasta ese mes en el momento en el que fueron enviados los datos.

### **6.2.1. Análisis general 2023**

Para este análisis se crearon tres tablas, con sus respectivas gráficas, siendo clasificadas de la siguiente forma.

Producción total de Tucan 3 durante el año 2023, recopilando la producción total de cada mes. Esta gráfica en particular permite observar ciertos patrones de cosecha en términos generales.

Promedio de producción por bloque, en la que se buscaba realizar una comparativa entre el promedio que junto a la herramienta Python se calculó con la producción total de cada uno de los bloques, permitiendo ver un comportamiento en dimensión general y otro más específico respectivamente, permitiendo igualmente identificar si hay uniformidad en los datos.

Por último, la producción total por variedad, en la que se puede identificar las variedades con mayor y menor producción.

### 6.2.2. Análisis descriptivo de cada bloque 2023

La tabla presentada a continuación evidencia la cantidad de filas, la media, la desviación estándar, el valor mínimo, el valor máximo y porcentajes de ubicación de los datos que servirán posteriormente para la realización de un gráfico de cajas y bigotes.

**Tabla 6**

*Estadística descriptiva del bloque 1 en el año 2023*

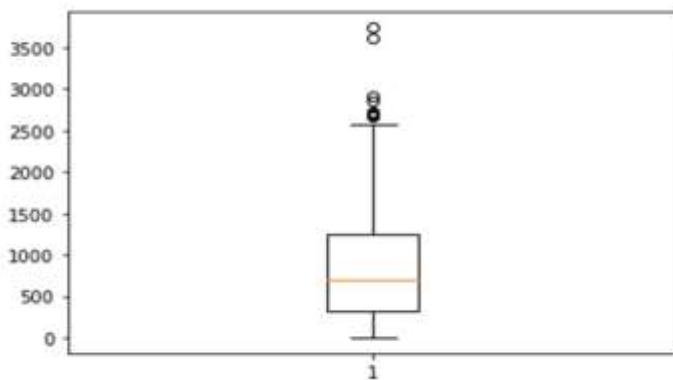
	BLOQUE	DÍA	SEMANA	MES	VARIEDAD	PRODUCCIÓN
Número de filas	648	648	648	648	648	648
Media	1	15,43519	15,93056			817,1635802
Desviación estandar	0	8,847259	8,909633			663,9452952
Valor Mínimo	1	1	1			0
25%	1	8	8			313
50%	1	15	16			708
75%	1	23	24			1246,25
Valor Máximo	1	31	31			3743

Fuente: elaboración propia

Posteriormente se realizó la gráfica de Cajas y Bigotes estándar que permite observar que tan variables son los datos en promedio. Este último gráfico, junto con la tabla anterior permite identificar que el promedio de producción, en el caso del Bloque 1 es de 817 unidades, un supuesto máximo de 3.743, teniendo un comportamiento entre 313 y 1.246 unidades.

**Gráfico 4**

*Diagrama de cajas y bigotes para el Bloque 1 de Tucan 3 durante el año 2023*



Fuente: elaboración propia

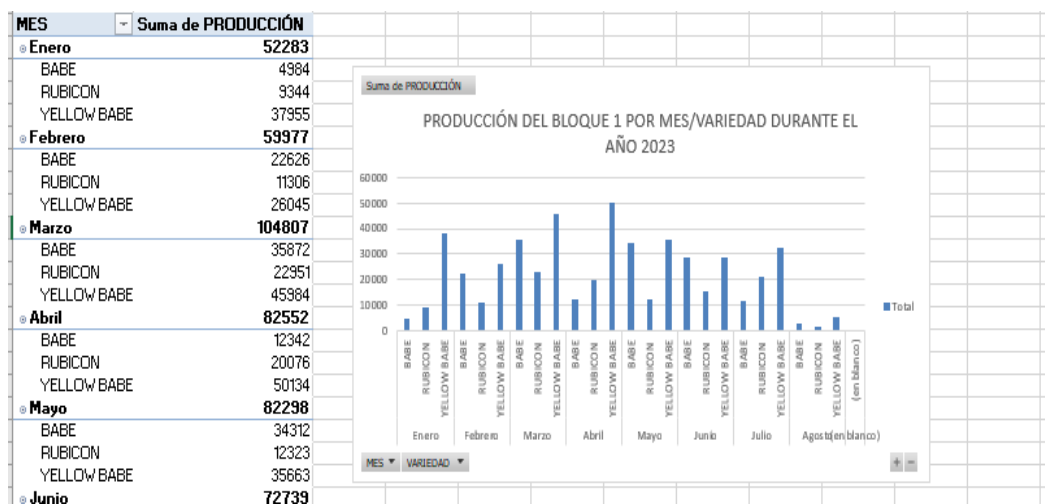
De la misma forma en la que se realizó el anterior procedimiento, se hizo lo mismo con el resto de los bloques, pero para evitar congestión de información, se dispuso un anexo llamado anexo 3 con el fin de ampliar la información.

### 6.2.3. *Análisis específico 2023*

En cada bloque, se dispuso con los datos obtenidos del ejercicio de programación una tabla y un gráfico dinámicos, herramientas que permiten interacción con los datos de forma visual y numérica, sin necesidad de alterarlos de forma radical. Esto indica que si alguna persona involucrada en el proceso de producción desea ver, por ejemplo un comportamiento específico de una variedad en un mes, semana y/o día en particular.

### **Gráfico 5**

*Organización del análisis del bloque 1 de la producción de Tucan 3 para el 2022*



Fuente: elaboración propia

### 6.3. Reunión concluyente con el agrónomo

El 5 de diciembre de 2023, se llevó a cabo una reunión sustantiva con el agrónomo, durante la cual se presentó de manera integral y detallada los avances y resultados obtenidos a lo largo del semestre. La estructura de la exposición abarcó diversos aspectos cruciales, explorando exhaustivamente los objetivos que se propusieron en relación con el proyecto para Tucan Flowers S.A. Esto incluyó metas específicas y medibles que se buscaban alcanzar en términos de eficiencia, calidad y optimización de procesos. A su vez, se presentaron los datos proporcionados por Tucan, detallando su relevancia y aplicación en el contexto del análisis del equipo. Este segmento sirvió como base fundamental para la investigación y posteriores conclusiones. Se detalló durante la reunión la metodología que se empleó para organizar y estructurar los datos facilitados. Este paso fue esencial para garantizar una manipulación efectiva y un análisis preciso de la información.

Se destacó, como parte fundamental la elección de Python como la herramienta de programación principal para procesar los datos. Esta elección se basó en su versatilidad, eficiencia y capacidad para manejar conjuntos de datos complejos, permitiendo igualmente la presentación de manera detallada los resultados obtenidos a través del análisis realizado por el equipo. Esto incluyó tendencias identificadas, patrones relevantes y cualquier hallazgo significativo que emergió de la interpretación de los datos.

### ***6.3.1. Recomendaciones Empresariales***

El equipo formuló recomendaciones específicas destinadas a optimizar los procesos de Tucan, basadas en las conclusiones extraídas del análisis final. Estas sugerencias se diseñaron para impulsar la eficiencia y el rendimiento general de la empresa. Asimismo, se plantearon preguntas estratégicas destinadas a obtener una comprensión más profunda de los resultados presentados. Estas preguntas fueron diseñadas para fomentar la discusión y proporcionar insights adicionales que podrían ser cruciales para la toma de decisiones.

La presentación fue recibida con una evaluación positiva por parte del agrónomo, quien destacó diversos aspectos clave. En particular, señaló que los cultivos en los bloques 1 y 2 exhiben un desarrollo más robusto y una mayor estabilidad, lo que se traducía en niveles de producción significativamente superiores en comparación con otros bloques. Además, resaltó la influencia determinante de factores externos, como el clima y las condiciones del suelo, en los niveles de producción. Destacó que, sin embargo, el éxito de los lotes estaba intrínsecamente ligado al manejo preciso de los tallos por parte del personal agrícola.

La observación del agrónomo amplió la comprensión de la complejidad inherente a la producción agrícola y subrayó la importancia de abordar múltiples variables para garantizar un rendimiento óptimo. Su énfasis en la gestión de los tallos como un factor crítico resalta la necesidad de prácticas agrícolas cuidadosas y específicas para maximizar la productividad.

Además, el agrónomo elogió la utilidad de las gráficas elaboradas en Excel, destacando que estas representaciones visuales proporcionaban un panorama práctico y fácilmente comprensible de la producción por bloques y variedad. Esta herramienta visual no solo facilita la interpretación de datos complejos, sino que también le permite al agrónomo tomar decisiones de manera más rápida y eficiente. La capacidad de obtener una visión general clara de la producción facilita la identificación de tendencias y la implementación ágil de estrategias para mejorar la eficiencia operativa.

## **7. Conclusiones**

A lo largo del año 2023, la colaboración del equipo con Tucan Flowers S.A. proporcionó una perspectiva valiosa sobre el funcionamiento interno de una empresa dedicada al cultivo de mini rosas. Este proyecto no solo permitió profundizar en los entresijos de la operación de la compañía, sino que también facilitó la identificación de manera precisa tanto sus fortalezas como sus áreas de oportunidad.

Un hallazgo fundamental que emerge de la inmersión hecha por el equipo en este proyecto es la trascendental importancia de la comunicación efectiva entre todas las áreas de

la empresa. Observamos cómo una coordinación fluida y un intercambio de información eficiente son elementos cruciales para optimizar la producción interna.

La sinergia entre departamentos se revela como un factor determinante para la eficacia operativa, desde la planificación hasta la ejecución de tareas cotidianas.

Asimismo, la gestión exitosa de una empresa agrícola como Tucan Flowers S.A. está intrínsecamente vinculada a la habilidad de mitigar riesgos y manejar de manera eficaz los factores externos. La variabilidad climática y las condiciones del suelo, por ejemplo, se erigen como variables que deben ser cuidadosamente consideradas y abordadas para garantizar una producción consistente y de calidad. Este enfoque en la gestión integral de riesgos destaca la necesidad de estrategias anticipativas y resilientes.

La importancia crítica de la automatización y el empleo de herramientas de programación en el contexto agrícola fue de igual modo uno de los factores a destacar durante el desarrollo del proyecto, ya que se identificó que la optimización de procesos a través de la automatización no solo aumenta la eficiencia operativa, sino que también reduce la posibilidad de errores humanos. Las Naciones Unidas, por su parte manifestó por medio de uno de los objetivos de desarrollo sostenible la necesidad de un cambio profundo en los sistemas de agricultura, priorizando tanto el avance tecnológico como el desarrollo de herramientas para que los trabajadores del sector tengan la capacidad de mejorar y posteriormente trabajar sus tierras en pro a un beneficio vertical y sostenible. (United Nations, s.f.). Los factores mencionados con anterioridad evidencian que implementación de herramientas de programación, como el uso de Python en el caso del proyecto reveló como un catalizador clave para facilitar la toma de decisiones informadas y ágiles, pero que dejan de ser integrales si no se tiene en cuenta el factor de mano de obra y el trabajo realizado por los encargados de tratar y producir los cultivos, quienes terminan siendo un pilar fundamental para cualquier compañía de carácter agroindustrial.

## 8. Bibliografía

- Alexandre. (2022, diciembre 19). Pandas: La biblioteca de Python dedicada a la Data Science. *Formation Data Science | DataScientest.com*. <https://datascientest.com/es/pandas-python>
- Chaves, S. (2022, febrero 21). El lenguaje Python y sus principales librerías. *Formadores IT*. <https://formadoresit.es/el-lenguaje-python-y-sus-principales-librerias/>
- Izary Rondón. (2022, noviembre 2). ▷ *Tipos de datos de Python | EIP*. <https://eiposgrados.com/blog-python/tipos-de-datos-de-python/>
- Ministerio de Agricultura. (2021). *Cadena de Flores, Follajes y ornamentas*. Ministerio De Agricultura. <https://sioc.minagricultura.gov.co/Flores/Documentos/2021-03-31%20Cifras%20Sectoriales.pdf>
- Perakis, K., Lampathaki, F., Nikas, K., Georgiou, Y., Marko, O., Maselyne, J. (2020). *CYBELE – Fostering precision agriculture & livestock farming through secure access to large-scale HPC enabled virtual industrial experimentation environments fostering scalable big data analytics*. Science Direct. <https://www.sciencedirect.com/science/article/pii/S1389128619305353>
- Pursell, S. (s. f.). *Guía completa para el análisis de datos (con ejemplos)*. Recuperado 7 de diciembre de 2023, de <https://blog.hubspot.es/marketing/analisis-de-datos>
- Statista (2023, octubre 16). *Ocupación en el sector agrario y pesquero en Colombia 2022*. Statista Research Department. <https://es.statista.com/estadisticas/1337185/ocupacion-en-agricultura-ganaderia-y-pesca-colombia>
- T. Rehman, C. Romero, (1993) *The application of the MCDM paradigm to the management of agricultural systems: some basic considerations*. Science Direct. <https://www.sciencedirect.com/science/article/abs/pii/0308521X9390002J>

Team, D. (2022, mayo 27). *¿Qué es un DataFrame? -. Formation Data Science / DataScientest.com*. <https://datascientest.com/es/que-es-un-dataframe>

United Nations (s.f.), *Goal 2: Zero Hunger. Sustain.* <https://www.un.org/sustainabledevelopment/hunger/>.