



Segmentación de operarios de raleo de uva Campaña 21 – PIURA

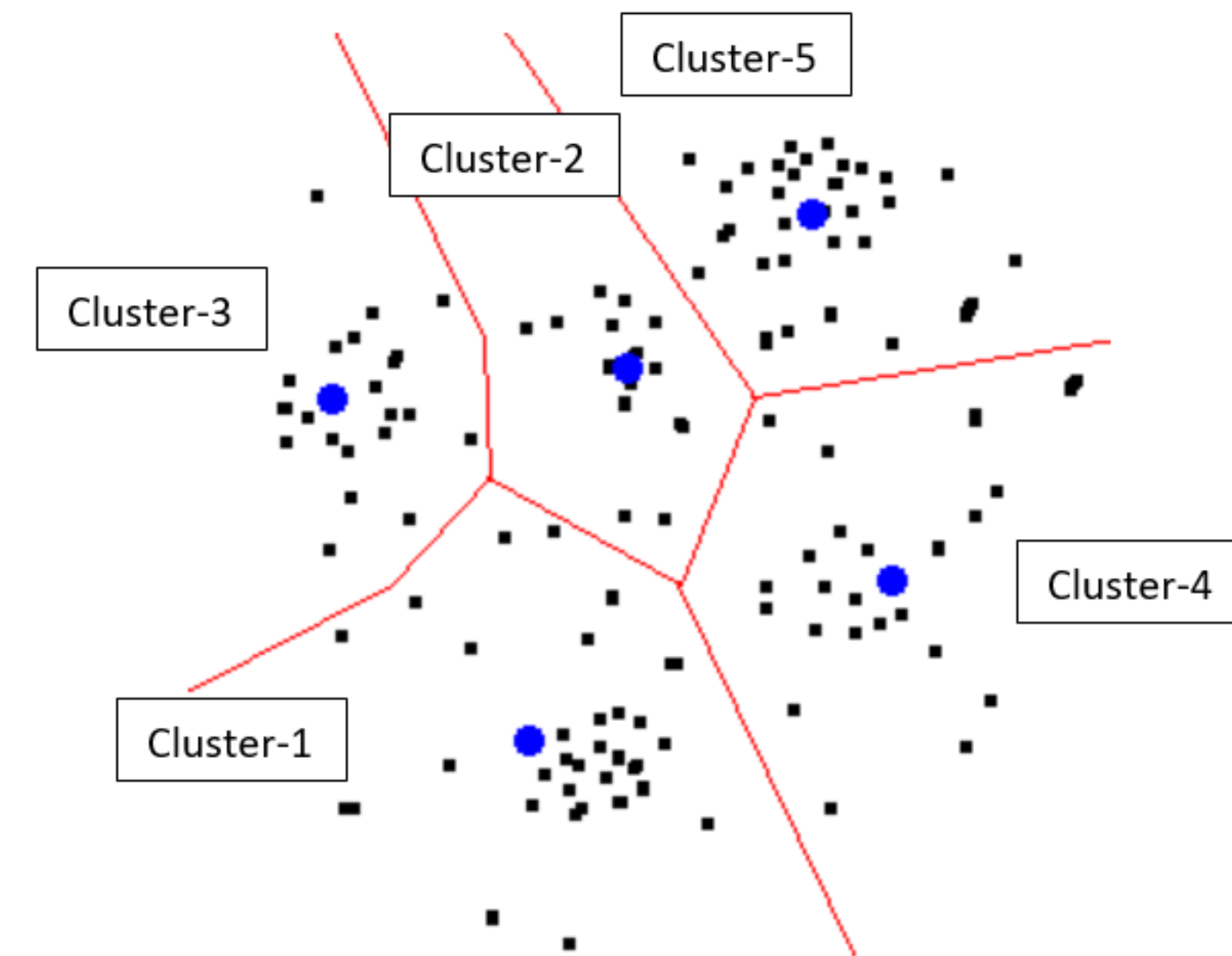
Pedregal
Noviembre 2021

Segmentación

(Análisis clúster)

Técnica estadística multivariada que tiene como objetivo la clasificación o agrupamiento de individuos u objetos en clases o conglomerados a partir de mediciones realizados en ellos, basado en dos principios fundamentales:

- ❑ Que cada grupo sea homogéneo respecto a las variables utilizadas para caracterizarlos; es decir, que cada observación sea **parecida** a todas las que están incluidas en ese grupo.
- ❑ Que los grupos sean **lo más distintos posible** unos de otros respecto a las variables consideradas.



Algoritmo k-means

(Mc Queen, 1967)

1. Especificar el número k de clusters que se desea crear
 2. Seleccionar de forma aleatoria k observaciones del conjunto de datos como centroides iniciales. Centroide: media de las variables que conforman el clúster
 3. Reubicar cada una de las observaciones al centroide más cercano
 4. Para cada uno de los k-clusters recalcular y actualizar su centroide
 5. Repetir los pasos 3 y 4 hasta que se satisfaga un criterio de convergencia (las reubicaciones no cambien o se alcance el número de iteraciones establecido)
- ✓ Métodos para definir el número k óptimo de clusters: Elbow method, Average silhouette method, Gap statistics method
 - ✓ Validación de los clusters: Algoritmo de clasificación (Análisis Discriminante) y descripción de los clusters

Medidas de desempeño del algoritmo:

Suma de cuadrados entre clusters

Error de clasificación

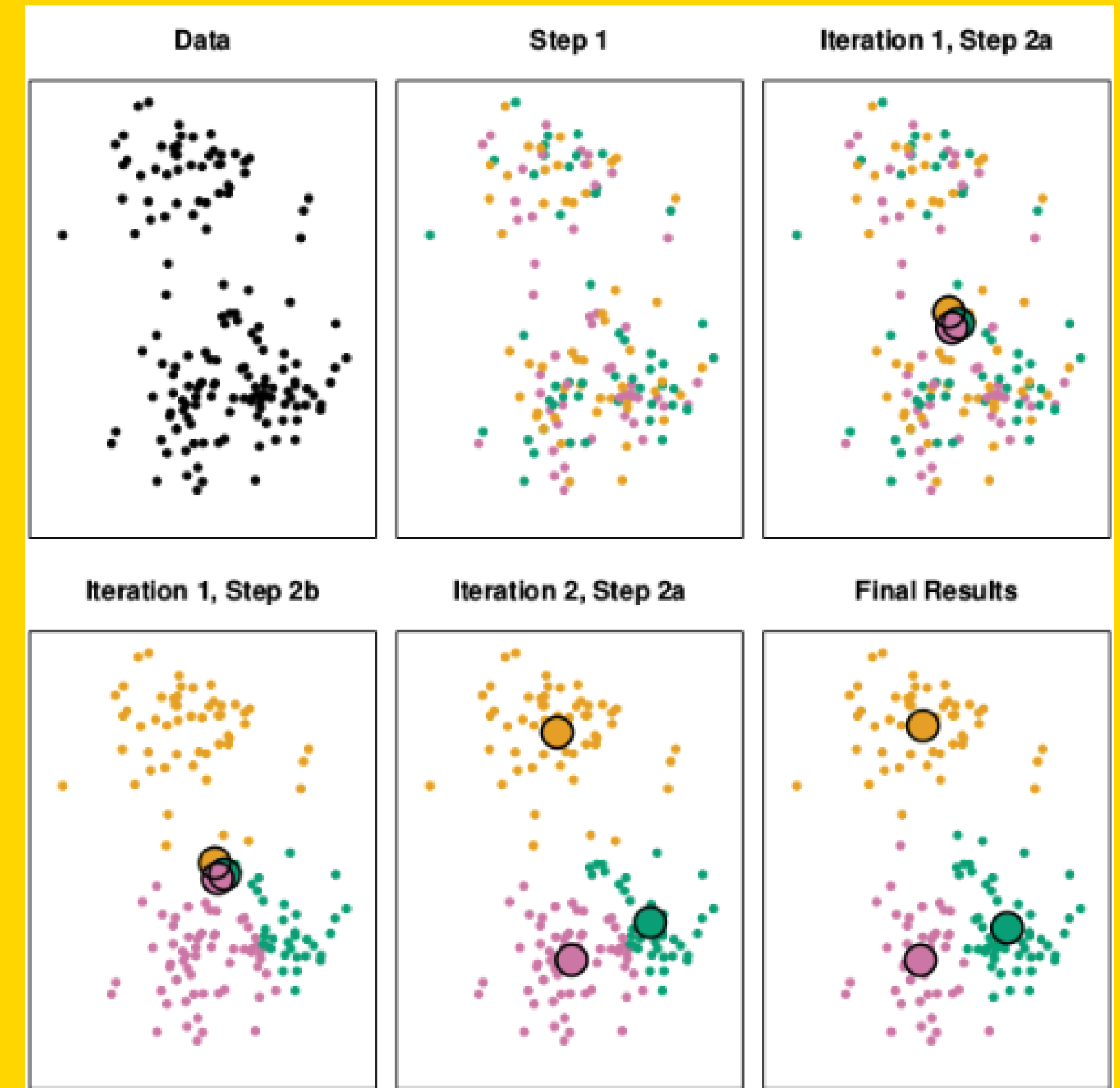


Gráfico extraído de:
https://www.cienciadedatos.net/documentos/37_clustering_y_heatmaps
#K-means_clustering



¿Qué perfiles de operarios de raleo encontramos en las campañas 19 y 20?

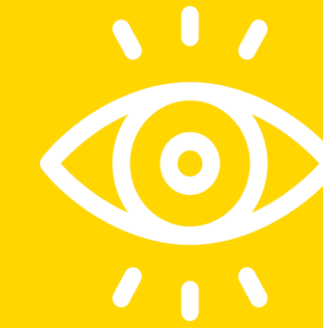
Encontramos **4 segmentos**

$$\text{Productividad (racimos/hora)} = \frac{\text{Total racimos por persona}}{\text{Total de horas laboradas por persona}}$$

En las campañas 19 y 20 tanto en Ica como en Piura, el cluster productivo de raleo estuvo conformado por el 42% del total de operarios y en conjunto realizaron el 86% de la producción. Los jóvenes de baja productividad realizaron el 14% de la producción y representaron el 57% del total de operarios.

Cluster productivo

RALEO PEDREGAL	Mujeres jóvenes altamente productivas (14%)	Mujeres jóvenes productivas (27%)	Hombres adultos leales y responsables (2%) 42% ↓	Jóvenes de baja productividad (57%)
Productividad	70,6	54,2	58,5	31,8
% producción total en últimas 3 campañas	48%	36%	2% 86%	14%
Características demográficas	Principalmente mujeres. Tienen 31 años en promedio. Solteras y con 1 hijo. Tienen más operarios iletrados, con primaria incompleta/completa y secundaria incompleta que el resto	Predominan mujeres. Tienen 31 años en promedio. Solteras y tienen 1 hijo. Tienen más gente con secundaria incompleta que las jóvenes altamente productivas	Predominantemente hombres. Tienen 40 años en promedio. No tienen hijos y contiene más operarios casados que el resto. Tienen secundaria completa y mayor educación superior.	Tienen 29 años en promedio. Solteros. Hay más operarios con secundaria incompleta que el resto.
Sobre Pedregal y la cosecha	Son los operarios más productivos y constantes que el resto. Tienen experiencia en la labor pero no todos, pues hay operarios que han trabajado en una sola campaña. Viven relativamente cerca al fundo. La rotación es menor.	Operarios medianamente productivos. Han trabajado en promedio 2 campañas. Viven alejadas del fundo que las altamente productivas	Operarios medianamente productivos. Son los más antiguos de Pedregal y más experimentados en la labor. Viven más cerca del fundo. Tiene alta rotación en la labor.	Son los operarios menos productivos e inconstantes. Han trabajado una sola campaña principalmente. Viven bastante alejados del fundo. Tienen mayor rotación
Sobre deudas en el sistemas financiero	Tiene negocios, menos deuda y su calificación es normal. Tienen menor atraso de pago que el resto	Tiene negocios. Menos deuda pero mayor atraso en el pago que las altamente productivas	Tienen más productos bancarios y más deuda, tienen buena calificación. No tienen negocios	Tienen negocios, menos deuda y mayor atraso de pago



En la campaña 21 y en Piura ¿Cómo se reconformaron estos perfiles?

Proceso de la segmentación

Los datos analizados corresponden a la campaña 21 de raleo

1

Recolección de la información

- 1) Matriz de operarios de raleo en Piura:
- Parte diario
 - ERM- personal (declarado)
 - Nómina
- 3) RCC – Reporte de deudas Superintendencia de Banca y Seguros
- 2) RENIEC

2

Filtros de segmentación

El universo estuvo conformado por operarios de raleo de la campaña 21 – Piura

3

Variables de segmentación

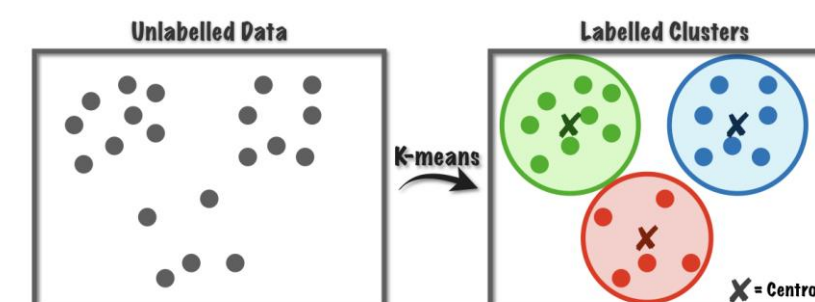
Identificamos variables del operario que discriminaron las diferencias entre operarios:

- Productividad
- Importe fijo
- Antigüedad
- Sexo
- Edad
- Número de hijos
- Indicador de pobreza
- Cantidad de campañas trabajadas

4

Análisis cluster

Aplicamos el algoritmo de clustering basado en centroides: K- means



5

Perfil de los clusters

Se obtuvieron 4 clusters de operarios de raleo y se perfilaron según:

- Sexo
- Edad
- Grado de instrucción
- Estado civil
- Tenencia de hijos
- Antigüedad
- Distancia domicilio- Pedregal (*)
- Cantidad de campañas
- Importe (fijo, bonos, horas extras)
- Bancarización

* Distancia ortodrómica: Es el arco de círculo máximo que corresponde a la distancia más corta entre dos puntos del globo terrestre. La fuente fue RENIEC 2021

RALEO – PIURA

Operarios productivos (50%)↑

Campañas 19-20: 42%

	S1	S2	S3
Productividad	158,5	77,3	71,7

- S1. Hombres jóvenes altamente productivos y responsables (10%)
- S2. Jóvenes productivos con carga familiar (20%)
- S3. Operarios hombres con experiencia (20%)

S4. Jóvenes de baja productividad (50%)

Operarios menos productivos (50%)↓


Campañas 19-20: 57%

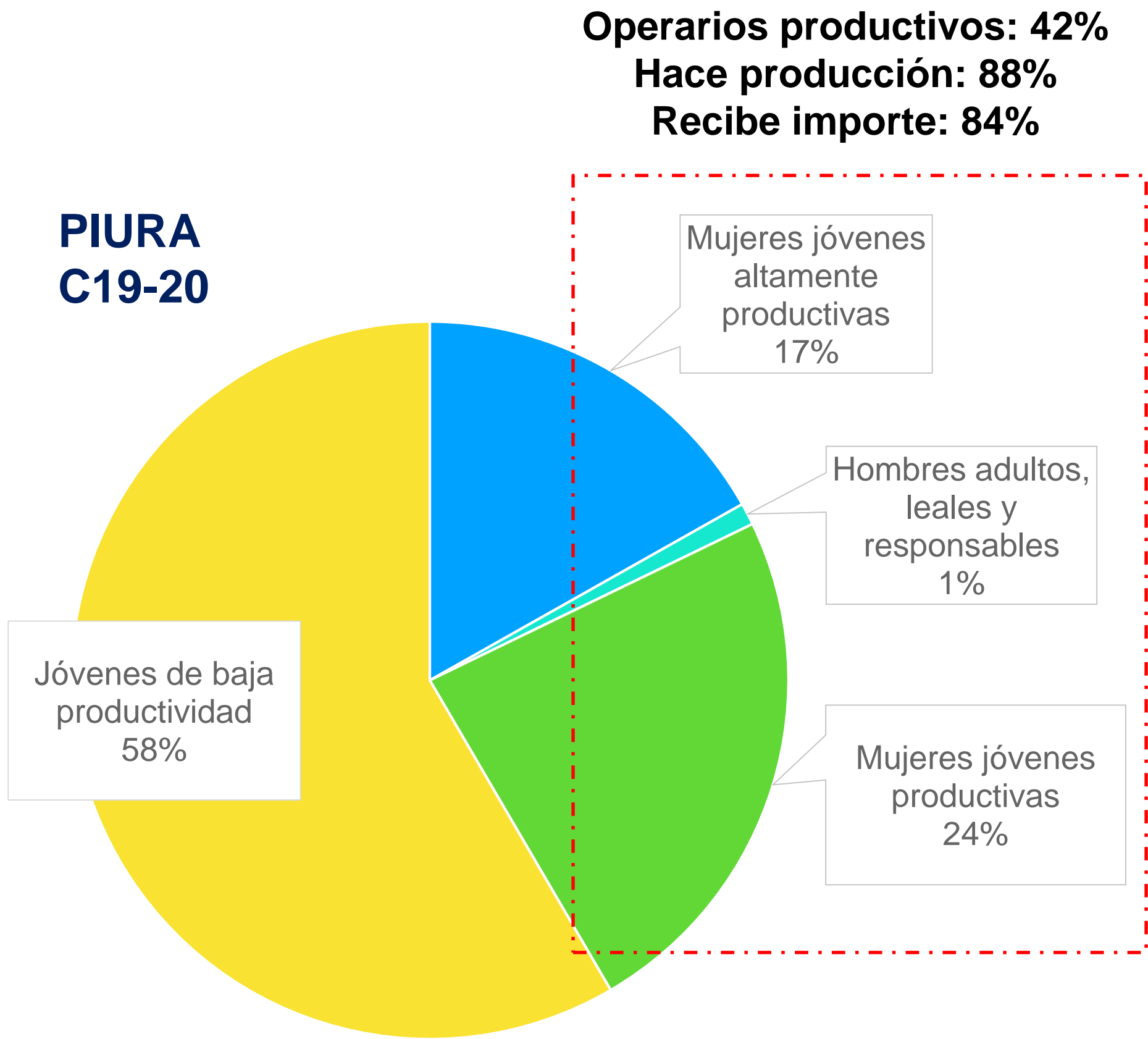
	S4
Productividad 21	32.0↑
Productividad 19-20	23,4

Algoritmo k-means: 4 clusters
Suma de cuadrados entre clusters = 89,3%
Error de clasificación = 2,7%
Software: R

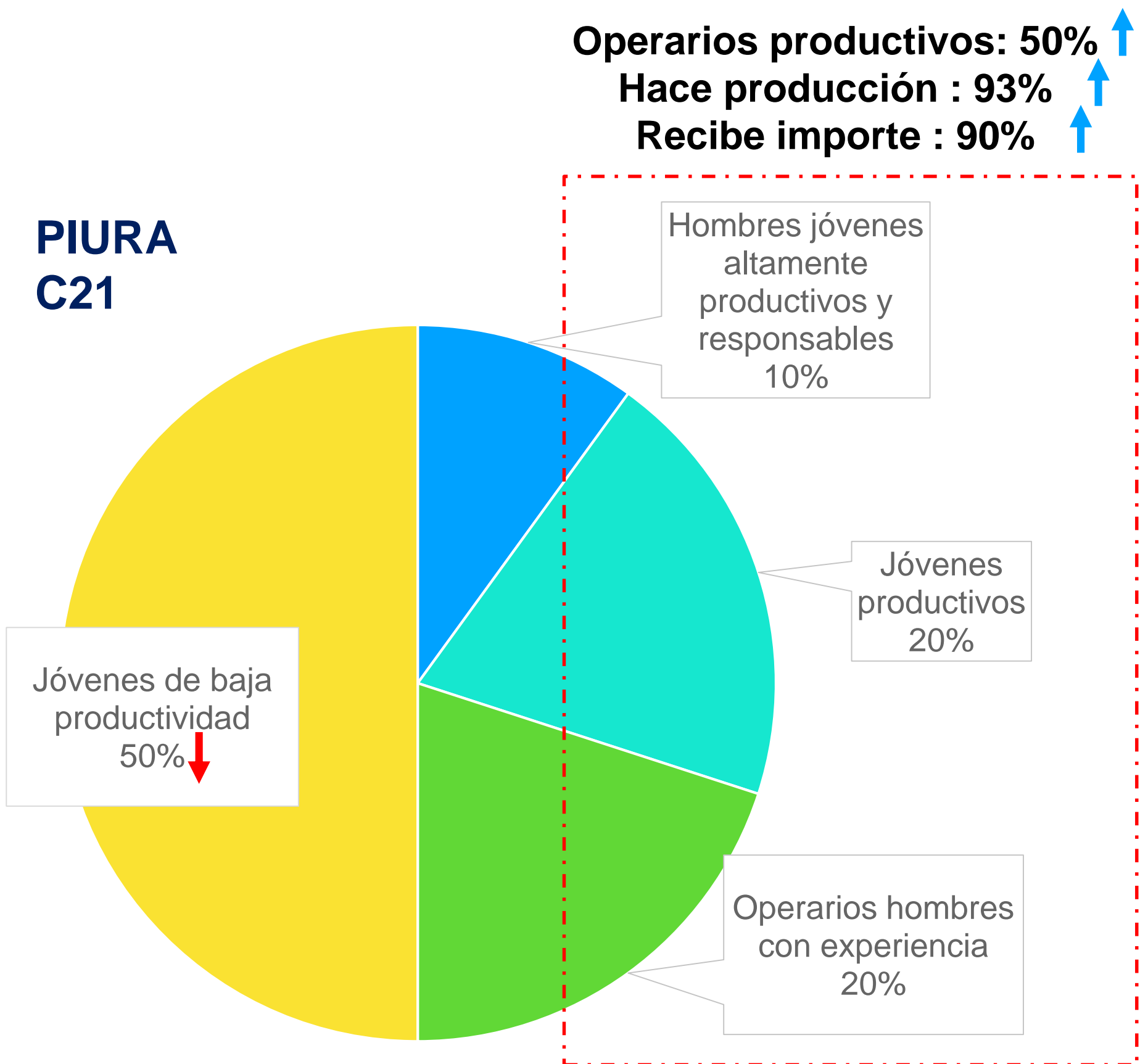
↑ Incremento significativo
↓ Disminución significativa

Para la campaña 21 se incrementa la productividad en el raleo y la presencia de jóvenes de baja productividad disminuyó de 58% a 50%. El cluster productivo se ha reconfigurado.



Piura – Campaña	19	20	21
Productividad raleo	41,3	30,5	57 



Base: 12,677 operarios de raleo en las campañas 18,19 y 20



Base: 3,447 operarios de raleo en la campaña 21

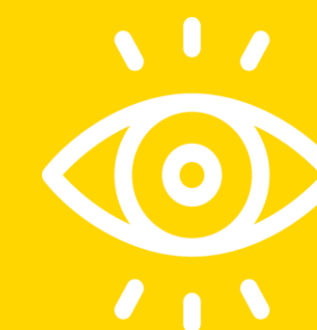
 Incremento significativo
 Disminución significativa



En la campaña 21 de raleo en Piura, los jóvenes de baja productividad representaron el 50% de los operarios y realizaron el 7% de la producción, así también recibieron el 10% del importe desembolsado para la labor. Por otro lado el 50% restante que conforma el grupo de operarios con mayor productividad realizó el 93% de la producción.

Raleo – PIURA	Hombres jóvenes altamente productivos y responsables	Jóvenes productivos	Operarios hombres con experiencia	Jóvenes de baja productividad	
	Productividad	158,5	77,3	71,7	32,0
	Producción total (racimos)	15'113,819 (42%)	11'502,507 (32%)	6'370,784 (18%)	2'606,801 (7%)
	Importe total	1'633,132 (38%)	1'436,349 (34%)	779,137 (18%)	416,483 (10%)
	Cantidad de operarios	333 (10%)	683 (20%)	687 (20%)	1,744 (50%)

¿Cómo son los perfiles de los operarios de raleo de la campaña 21 en Piura?



Encontramos **4 segmentos** conformado por operarios jóvenes principalmente

$$Productividad \text{ (racimos/hora)} = \frac{\text{Total racimos por persona}}{\text{Total de horas laboradas por persona}}$$

Hombres jóvenes altamente productivos y responsables | 10%

+ **Productividad:** 158,5 racimos/hora

Principalmente son hombres **(75%)**. Tienen 26 años en promedio, solteros **(87%)** y con hijos **(65%)** uno en promedio. Su grado de instrucción es primaria principalmente **(62%)** y contiene más operarios iletrados **(6%)**.

Tienen una antigüedad en Pedregal de **7,1 meses** y han trabajado de 2 campañas a más **(66%)** en promedio. Son los que **rotan menos** en la labor.

En la campaña 21 han trabajado en 22 lotes, 3 fundos y las variedades **SHEEGENE 02 TIMPSON™**, **JACKS SALUTE™** y **RED GLOBE** principalmente.

Viven cerca del fundo **(a 38 km)**.

Tienen **3k** de deuda y buena calificación.

	PIURA				
Campaña	Total días campaña	Días laborados			% de días laborados por campaña (media)
		Mínimo	Máximo	Media	
21	57	23	44	36,5	64%



$$Productividad\ (racimos/hora) = \frac{Total\ racimos\ por\ persona}{Total\ de\ horas\ laboradas\ por\ persona}$$

Jóvenes productivos con carga familiar | 20%

Productividad: 77,3 racimos/hora

Contiene más mujeres que el resto de segmentos **(51%)**. Tienen 28 años en promedio, son solteros **(87%)** y tienen hijos **(65%)**. Tienen primaria y secundaria incompleta principalmente **(76%)**.

Tienen una antigüedad en la empresa de **5,7 meses** y han trabajado **2,1 campañas** en promedio.

En la campaña 21 han trabajado en 12 lotes, 3 fundos y las variedades **SHEEGENE 02 TIMPSON™, SHEEGENE 03 MAGENTA™ y JACKS SALUTE™** principalmente.

Viven relativamente cerca al fundo **(a 55 km)**.

Tienen un negocio **(49%)**. En caso de tener deuda, los días de atraso es menor al resto.



	PIURA				
Campaña	Total días campaña	Días laborados			% de días laborados por campaña (media)
		Mínimo	Máximo	Media	
21	57	11	36	24,7	43%

$$Productividad\ (racimos/hora) = \frac{Total\ racimos\ por\ persona}{Total\ de\ horas\ laboradas\ por\ persona}$$

Operarios hombres con experiencia | 20%

Productividad: 71,7 racimos/hora

Principalmente son hombres **(56%)**. Tienen 29 años en promedio pero también contiene operarios con más de 31 años **(45%)** y tienen hijos **(57%)**. Tienen secundaria completa e incompleta **(56%)**.

Comparado con el resto son las más antiguos, han trabajado en Pedregal por **8,3 meses** y han trabajado hasta **2 campañas en promedio**.

En la campaña 21 han trabajado en 6 lotes, 2 fundos y las variedades **SHEEGENE 02 TIMPSON™** y **RED GLOBE** principalmente.

Viven alejados del fundo **(a 152 km)**.

No tienen negocio **(60%)**.



	PIURA				
Campaña	Total días campaña	Días laborados			% de días laborados por campaña (media)
		Mínimo	Máximo	Media	
21	57	2	36	13,0	23%

Productividad (racimos/hora) = $\frac{\text{Total racimos por persona}}{\text{Total de horas laboradas por persona}}$

Jóvenes de baja productividad | 50%

- **Productividad:** 32,0 racimos/hora

Son operarios de 27 años en promedio y principalmente solteros comparado al resto de segmentos **(91%)** y sin hijos **(59%)**. Tienen secundaria como grado de instrucción **(62%)**.

Contiene a los operarios que han trabajado menos tiempo en la empresa: **4,1 meses en promedio** y una sola campaña principalmente **(69%)**. Este grupo es grande porque contiene a operarios con **mayor rotación**.

En la campaña 21 han trabajado en 2 lotes, 1 fundo y la variedad **SHEEGENE 02 TIMPSON™** principalmente.

Viven más alejados del fundo **(a 173 km)**.

Tienen un negocio **(50%)**. Los que tienen deuda presentan mayor atraso de pago que el resto **(44%)**.

	PIURA				
Campaña	Total días campaña	Días laborados			% de días laborados por campaña (media)
		Mínimo	Máximo	Media	
21	57	1	13	3,7	6%



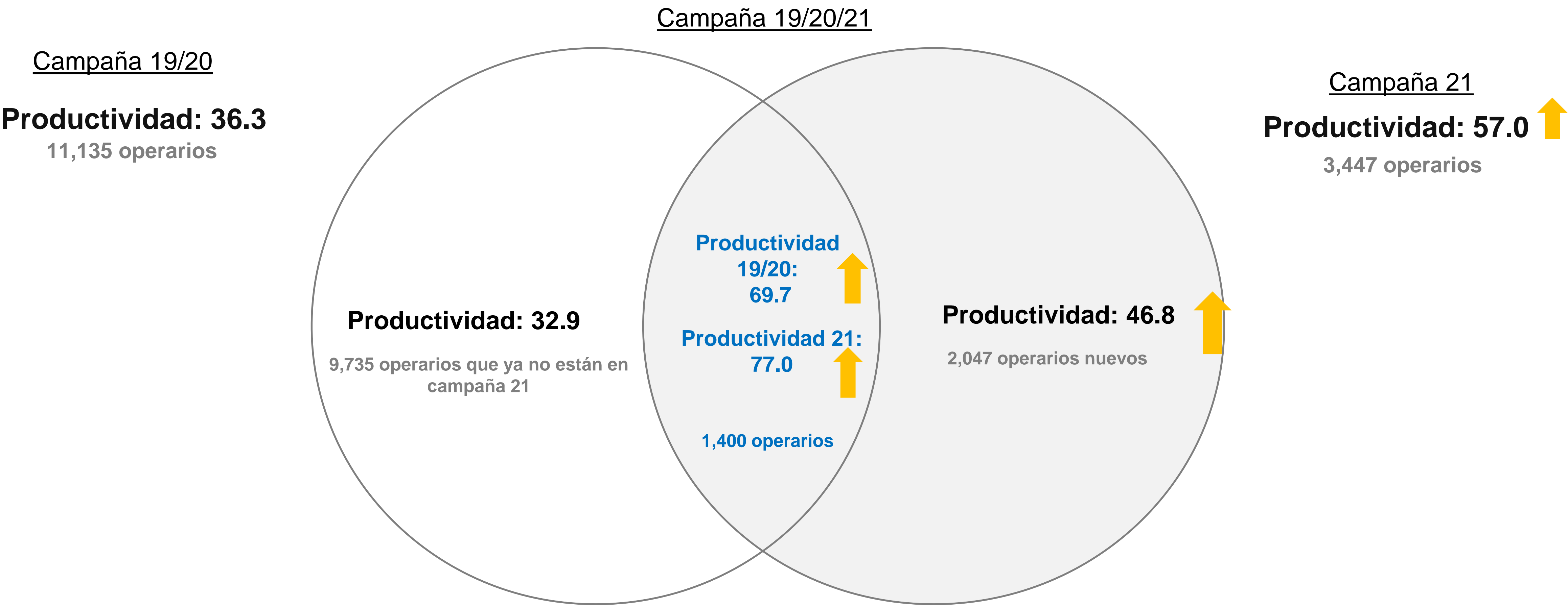
$$Productividad\ (racimos/hora) = \frac{Total\ racimos\ por\ persona}{Total\ de\ horas\ laboradas\ por\ persona}$$

Encontramos mayor productividad entre los operarios que trabajaron más días en la labor, por el contrario la menor productividad estuvo entre los jóvenes con mayor rotación. La distancia del domicilio del operario (*) al fundo tuvo una relación inversa con la productividad, mientras más cerca viven al fundo más productivos fueron.

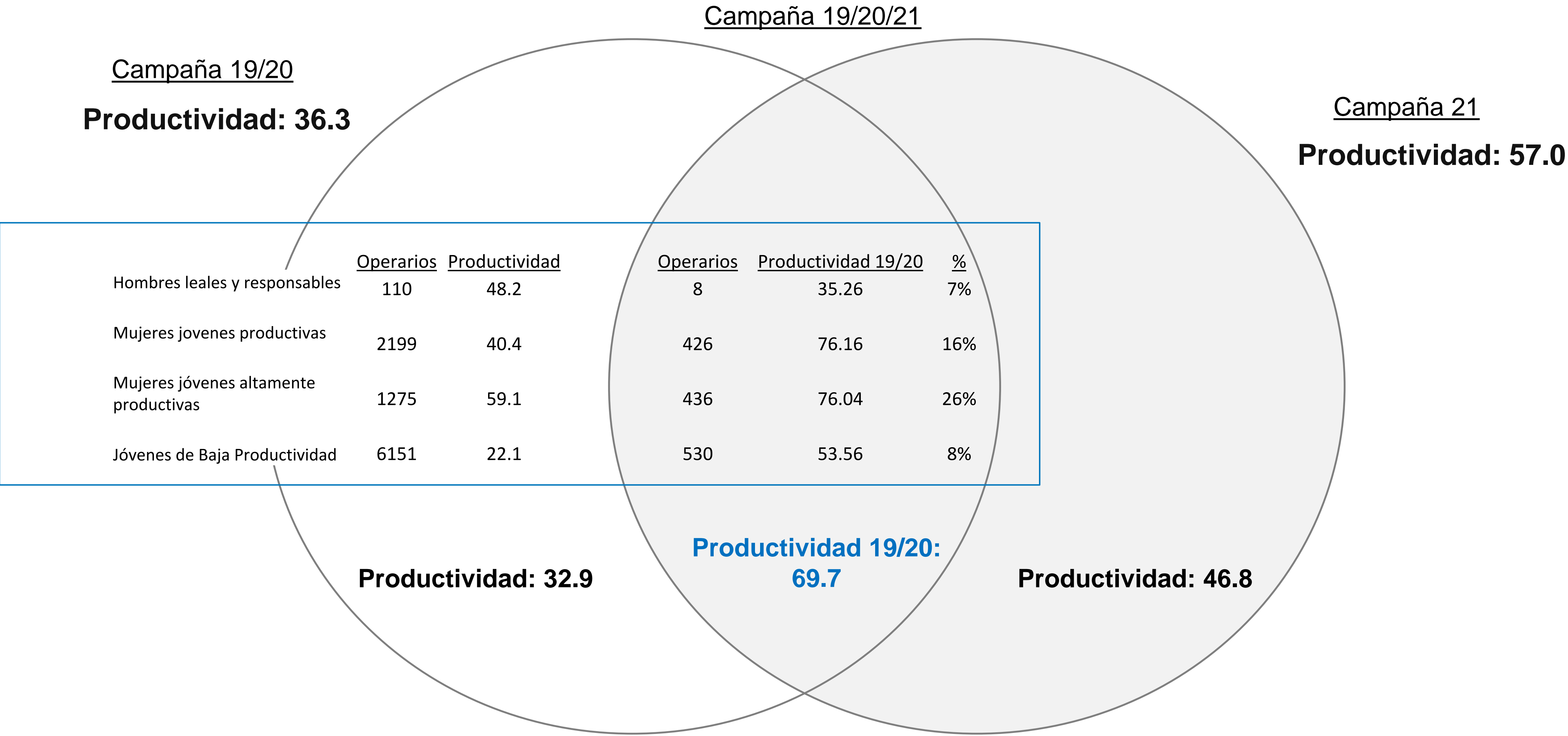
	Producción Total (%)	Importe Total (%)	Productivi dad	Total días campaña	Días laborados (media)	% de días laborados por campaña (media)	Distancia al fundo km (media)
Hombres jóvenes altamente productivos y responsables (10%)	42%	38%	158,5	57	36,5	64%	38 km
Jóvenes productivos (20%)	32%	34%	77,3	57	24,7	44%	55 km
Operarios hombres con experiencia (20%)	18%	18%	71,7	57	13,0	23%	152 km
Jóvenes de baja productividad (50%)	7%	10%	32,0	57	3,7	6%	173 km

(*) La distancia del domicilio al fundo fue estimado usando la distancia ortodrómica, es decir el arco de círculo máximo que corresponde a la distancia más corta entre dos puntos del globo terrestre. Las coordenadas del domicilio del operario fueron obtenidos de RENIEC 2021

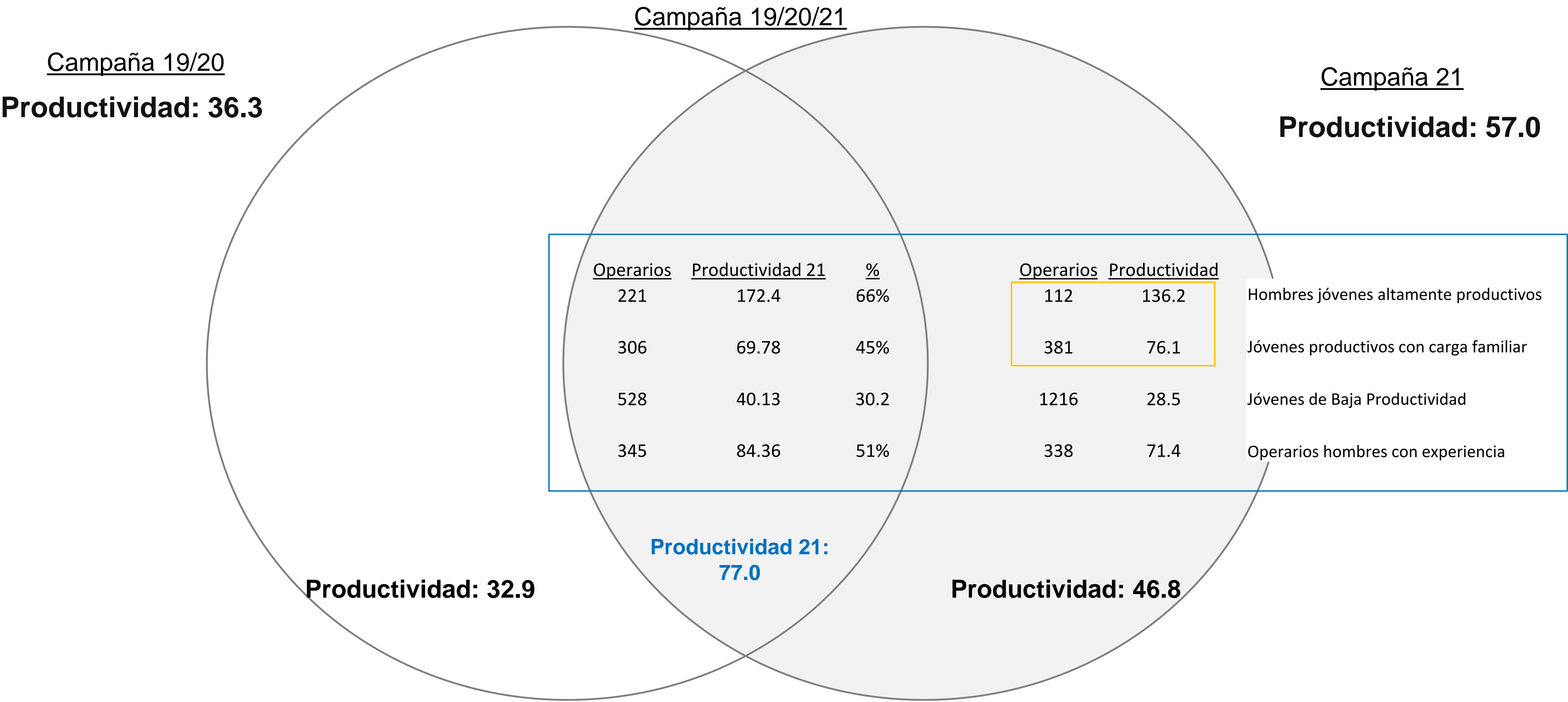
¿Como fue la productividad de los operarios de raleo que no se encuentran en la campaña 21, los que se mantuvieron y los nuevos?
La productividad de los operarios que continuaron y los nuevos fue mayor a los que dejaron la labor.



¿Cómo fue la distribución de los segmentos encontrados en la campañas 19/20?



¿Cómo fue la distribución de los segmentos encontrados en la campaña 21?



Conclusiones 2021:

- ✓ Los clúster no varían significativamente 2019, 2020 y 2021.
 - En 2021 se encontró un clúster de Joven de Hombres altamente productivo y otro clúster productivo donde se encuentran las mujeres.
 - Existe un clúster conformado por el 50% de operarios que presenta productividades muy inferiores al promedio y que aporta el 7% a la producción.
- ✓ Lo niveles productividad mejoran para el 2021:
 - 40% de los operarios en el 2021 esta compuesto por operarios del 2019 y 2020, con productividades de 77 racimos por hora (promedio 57).
 - El otro 60%, esta conformado por operarios nuevos, los perfiles de Jóvenes Productivos con carga familiar y Hombres Jóvenes Altamente Productivos incrementan el promedio de productividad en estos operarios nuevos. El perfil de nuevos de baja productividad en las 3 campañas (19/20/21) muestras productividades de 25.
- ✓ En 2021 los clúster productivos (50%) hacen el 93% de la producción, muy similar a los años 2019/2020.

Recomendaciones 2021:

- ✓ Acciones de retención (programa de ayuda para sus hijos) en el cluster jóvenes productivos con carga familiar.
- ✓ Incentivar la constancia del trabajador a través de bonos o rangos por antigüedad y productividad.
- ✓ Recolectar información cualitativa para entender a profundidad las motivaciones del operario que impulsan su productividad
- ✓ Trabajar en un algoritmo predictivo con información de las bases de datos (edad, sexo, etc) e información cualitativa recogida en campo. El algoritmo calculara la probabilidad de que un operario sea productivo en la próxima campaña, así poder enfocarse en los más productivos e incluso sirva como filtro de selección.
- ✓ Generar un brainstorming de acciones de negocio con los clúster, enfocados en mantener los clúster productivos

A large yellow circle is positioned in the top-left corner of the slide, partially overlapping the white background.

Construyamos una “Gran historia” juntos

Contacto

gaby@yellowbrain.pe

yellowbrain.pe