

Actividad 5

```
#Obtengo un análisis univariado de una variable categórica en específico
table2 = freq_tbl(Micro_Retailer['314_reception_delivery_packages'])
table2
```

	314_reception_delivery_packages	frequency	percentage	cumulative_perc
0	Totalmente en desacuerdo	46	0.269006	0.676471
1	Algo de acuerdo	9	0.052632	0.808824
2	Totalmente de acuerdo	8	0.046784	0.926471
3	Ni de acuerdo ni en desacuerdo	3	0.017544	0.970588
4	Algo en desacuerdo	2	0.011696	1.000000

se realizó un análisis univariado de la variable '314_reception_delivery_packages'. Permite comprender la distribución de respuestas sobre la recepción y entrega de paquetes, identificando claramente que una mayoría significativa está "totalmente en desacuerdo", mientras que solo un pequeño porcentaje está "algo en desacuerdo".

```
#Obtengo un filtro de los valores más relevantes de las variables categóricas seleccionadas más significativas
Filtro1 = table2[table2['frequency']>3]
Filtro1
```

	314_reception_delivery_packages	frequency	percentage	cumulative_perc
0	Totalmente en desacuerdo	46	0.269006	0.676471
1	Algo de acuerdo	9	0.052632	0.808824
2	Totalmente de acuerdo	8	0.046784	0.926471

se realizó un filtrado de los datos para seleccionar las categorías más relevantes basado en la frecuencia de las respuestas. se excluyen las categorías menores a 3, lo que puede ayudar a reducir el ruido en los datos y facilitar el análisis de las tendencias principales.

```
#Ajusto el índice de mi DataFrame
Filtro_index1=Filtro1.set_index('314_reception_delivery_packages')
Filtro_index1
```

	frequency	percentage	cumulative_perc
314_reception_delivery_packages			
Totalmente en desacuerdo	46	0.269006	0.676471
Algo de acuerdo	9	0.052632	0.808824
Totalmente de acuerdo	8	0.046784	0.926471

se ajustó el índice del DataFrame Filtro1 utilizando una de sus columnas para mejorar la organización y el análisis de los datos.

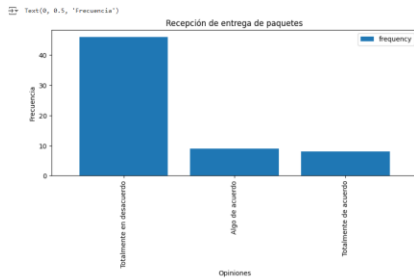
```
[ ] Filtro_index_frequency1= Filtro_index1.drop(['percentage', 'cumulative_perc'], axis=1)
Filtro_index_frequency1
```

	frequency
314_reception_delivery_packages	
Totalmente en desacuerdo	46
Algo de acuerdo	9
Totalmente de acuerdo	8

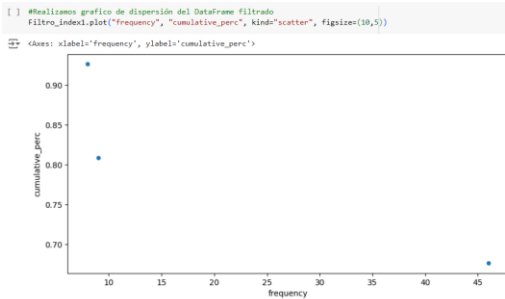
se realizó una operación para simplificar el DataFrame eliminando algunas columnas que no son necesarias. En este caso, mantener solo la columna de frecuencias permite un enfoque directo en la cantidad de respuestas para cada categoría.

```
#Realizamos gráfico de barras del dataframe filtrado
Filtro_index_frequency1.plot(kind='bar', width=0.8, figsize=(10,4))
plt.title('Recepción de entrega de paquetes')
plt.xlabel('Opiniones')
plt.ylabel('Frecuencia')
```

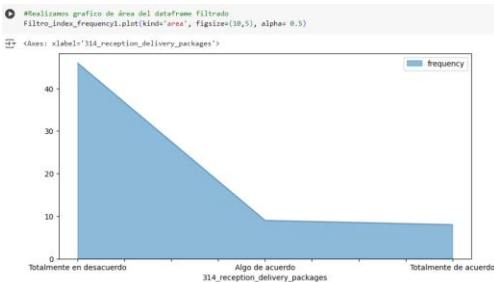
se realizó un gráfico de barras para visualizar las frecuencias de las respuestas categóricas sobre la recepción de entrega de paquetes.



El gráfico de barras proporciona las opiniones de los encuestados. La mayoría están "Totalmente en desacuerdo", mientras que solo una pequeña fracción está "Algo de acuerdo" o "Totalmente de acuerdo". Esto indica una fuerte tendencia hacia una perspectiva negativa sobre este aspecto.

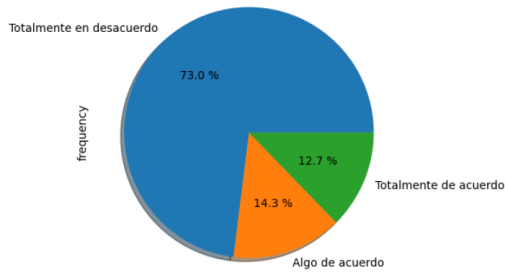


En el gráfico de dispersión, el punto más a la derecha, con la mayor frecuencia (46), tiene un porcentaje acumulativo más bajo (0.676471), lo que indica que la mayoría de las respuestas se concentran en esta categoría ("Totalmente en desacuerdo").



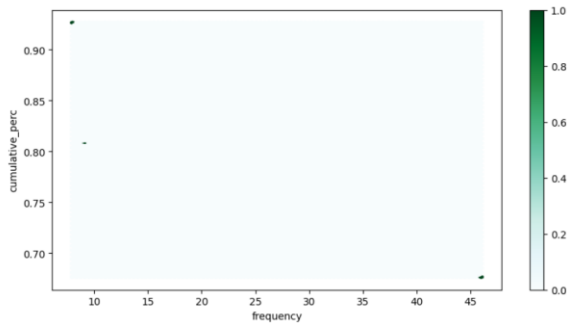
El gráfico de área es útil para mostrar de manera visualmente impactante la distribución de las respuestas categóricas en términos de frecuencia. El área debajo de la línea representa la frecuencia acumulada de cada categoría.

```
#Realizamos grafico de pastel del dataframe filtrado
Filtro_index1["frequency"].plot(kind='pie', figsize=(10,5), shadow=True, autopct="%0.1f %%")
<Axes: ylabel='frequency'>
```



El 73.0% de las respuestas indica que la mayoría de los encuestados están completamente en desacuerdo con la situación presentada. Un 14.3%, esta categoría es la segunda más alta, mostrando que una parte moderada de los encuestados está parcialmente de acuerdo y el 12.7% de las respuestas indica que está de acuerdo, lo que es relativamente bajo en comparación con las otras dos categorías.

```
#Realizamos grafico hexagonal del dataframe filtrado, alternativo al scatter plot
Filtro_index1.plot.hexbin("frequency", "cumulative_perc", figsize=(10,5))
<Axes: xlabel='frequency', ylabel='cumulative_perc'>
```



El gráfico hexagonal (hexbin) es generalmente útil para visualizar la densidad de grandes conjuntos de datos donde los puntos de datos se superponen. Sin embargo, en este caso específico con solo unos pocos puntos de datos, el gráfico se parece más a un gráfico de dispersión disperso.