In [ ]:
```
from src.main import generate_csv, read_csv, generate_summary, create_histogram, create_scatter_plot
import matplotlib.pyplot as plt
import os
```

In [ ]:
```

```

In [ ]:
```
default_file_folder = 'data'
csv_file_name = 'random_user.csv'
full_path = os.path.join(os.getcwd(),default_file_folder, csv_file_name)
params = {"results" : 500}
```

In [ ]:
```
file_path = generate_csv(file_name=full_path, params=params)
print(f'CSV generated at - {file_path}')
df = read_csv(file_path, engine_type='polars')
```

```
CSV data successfully saved to /Users/javidan/Developer/Data Engineering/Javidan_Polars_Descriptive_Statistics_Script/data/random_user.csv
CSV generated at - /Users/javidan/Developer/Data Engineering/Javidan_Polars_Descriptive_Statistics_Script/data/random_user.csv
```

Introduction to data

In [ ]:
```
df.head()
```

Out[ ]: shape: (5, 34)

gender	name.title	name.first	name.last	location.street.number	location.street.name	location.city	location.state	location.country	location.postcode	location.coordir
str	str	str	str	i64	str	str	str	str	str	
"male"	"Mr"	"Rahman"	"Stout"	5542	"Beun"	"Oene"	"Utrecht"	"Netherlands"	"1229 FC"	
"female"	"Mrs"	"Milica"	"Dinčić"	6782	"Mihajla Markovića "	"Batočina"	"Šumadija"	"Serbia"	"26293"	
"female"	"Miss"	"Nerea"	"Parra"	2877	"Avenida de América"	"Pontevedra"	"Aragón"	"Spain"	"23823"	
"female"	"Madame"	"Anne"	"Rousseau"	7867	"Rue du Moulin"	"Bözen"	"Ticino"	"Switzerland"	"8025"	
"female"	"Mademoiselle"	"Sonia"	"Gautier"	1216	"Place de L'Abbé-Georges-Hénocq…	"Herrliberg"	"Thurgau"	"Switzerland"	"9042"	

In [ ]:
```
df.shape
```

Out[ ]: (500, 34)

In [ ]:
```
res = generate_summary(df)
print(res)
```

```
shape: (9, 35)
```

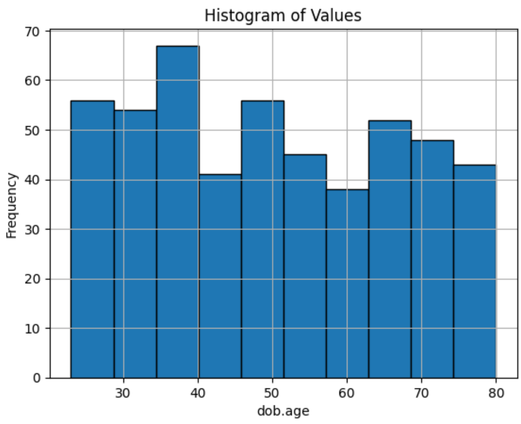
statistic --- str	gender --- str	name.title --- str	name.first --- str	…	picture.large --- str	picture.medium --- str	picture.thumbnail --- str	nat --- str
count	500	500	500	…	500	500	500	500
null_count	0	0	0	…	0	0	0	0
mean	null	null	null	…	null	null	null	null
std	null	null	null	…	null	null	null	null
min	female	Madame	Aada	…	https://randomuser.me/api/port…	https://randomuser.me/api/port…	https://randomuser.me/api/port…	AU
25%	null	null	null	…	null	null	null	null
50%	null	null	null	…	null	null	null	null
75%	null	null	null	…	null	null	null	null
max	male	Ms	یاسمین	…	https://randomuser.me/api/port…	https://randomuser.me/api/port…	https://randomuser.me/api/port…	US

In [ ]:
```

```

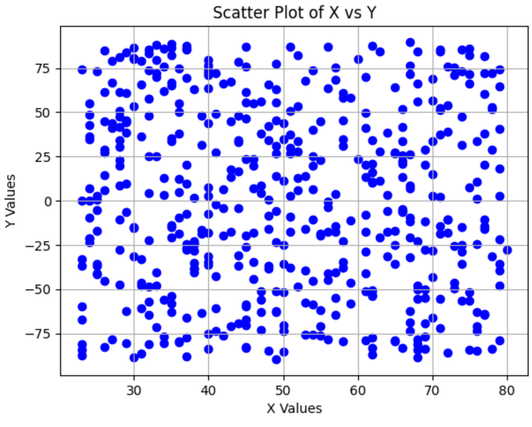
Visualization

In [ ]:
```
create_histogram(df, column = 'dob.age' )
```



Out[ ]: <module 'matplotlib.pyplot' from '/Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages/matplotlib/pyplot.py'>

In [ ]:
```
create_scatter_plot(df, x_col = 'dob.age', y_col='location.coordinates.latitude')
```




```
Scatter plot saved as 'scatter_plot.png'.
```

Out[ ]: <module 'matplotlib.pyplot' from '/Library/Frameworks/Python.framework/Versions/3.12/lib/python3.12/site-packages/matplotlib/pyplot.py'>

In [ ]:
```

```

In [ ]:
```

```