

Mini-Project Week 3

Output

Step 1: Import packages

```
import polars as pl
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

Step 2: Load dataset

```
df = pl.read_csv("gss2.csv", infer_schema_length=10000)
```

Step 3: Describe data frame

```
print(df.median())
print(df.describe())
```

shape: (1, 12)

region	income	happy	age	...	wrkstat	partyid	sex	year
---	---	---	---	---	---	---	---	---
str	str	str	str	...	str	str	str	f64
null	null	null	null	...	null	null	null	1993.0

shape: (9, 13)

describe	region	income	happy	...	wrkstat	partyid	sex	year
---	---	---	---	---	---	---	---	---
str	str	str	str	...	str	str	str	f64
count	53474	53474	53474	...	53474	53474	53474	53474.0
null_count	0	0	0	...	0	0	0	0.0
mean	null	null	null	...	null	null	null	1991.954894
std	null	null	null	...	null	null	null	10.948351
min	E. NOR. CENTRAL	\$1000 TO 2999	DK	...	KEEPING HOUSE	IND NEAR DEM	FEMALE	1973.0
25%	null	null	null	...	null	null	null	1983.0
50%	null	null	null	...	null	null	null	1993.0
75%	null	null	null	...	null	null	null	2002.0
max	W. SOU. CENTRAL	REFUSED	VERY HAPPY	...	WORKING PARTTIME	STRONG REPUBLICAN	MALE	2010.0

Step 4: Visualization

