```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

data = pd.read_csv('adult.csv')

data
```

	age	workclass	fnlwgt	education	education.num	marital.status	occupation	relationship	race	sex	capital.gain	capital.l
0	90	?	77053	HS-grad	9	Widowed	?	Not-in-family	White	Female	0	4
1	82	Private	132870	HS-grad	9	Widowed	Exec-managerial	Not-in-family	White	Female	0	4
2	66	?	186061	Some-college	10	Widowed	?	Unmarried	Black	Female	0	4
3	54	Private	140359	7th-8th	4	Divorced	Machine-op-inspct	Unmarried	White	Female	0	3
4	41	Private	264663	Some-college	10	Separated	Prof-specialty	Own-child	White	Female	0	3
...	...	...	...	...	...	...	...	...	...	...	...	
32556	22	Private	310152	Some-college	10	Never-married	Protective-serv	Not-in-family	White	Male	0	
32557	27	Private	257302	Assoc-acdm	12	Married-civ-spouse	Tech-support	Wife	White	Female	0	
32558	40	Private	154374	HS-grad	9	Married-civ-spouse	Machine-op-inspct	Husband	White	Male	0	
32559	58	Private	151910	HS-grad	9	Widowed	Adm-clerical	Unmarried	White	Female	0	
32560	22	Private	201490	HS-grad	9	Never-married	Adm-clerical	Own-child	White	Male	0	

32561 rows × 15 columns

Next steps: Generate code with `data` View recommended plots

```
data.head(4)
```

	age	workclass	fnlwgt	education	education.num	marital.status	occupation	relationship	race	sex	capital.gain	capital.loss
0	90	?	77053	HS-grad	9	Widowed	?	Not-in-family	White	Female	0	4356
1	82	Private	132870	HS-grad	9	Widowed	Exec-managerial	Not-in-family	White	Female	0	4356
2	66	?	186061	Some-college	10	Widowed	?	Unmarried	Black	Female	0	4356
3	54	Private	140359	7th-8th	4	Divorced	Machine-op-inspct	Unmarried	White	Female	0	3900

Next steps: Generate code with `data` View recommended plots

```
df = pd.DataFrame(data)
df
```

	age	workclass	fnlwgt	education	education.num	marital.status	occupation	relationship	race	sex	capital.gain	capital.l
0	90	?	77053	HS-grad	9	Widowed	?	Not-in-family	White	Female	0	4
1	82	Private	132870	HS-grad	9	Widowed	Exec-managerial	Not-in-family	White	Female	0	4
2	66	?	186061	Some-college	10	Widowed	?	Unmarried	Black	Female	0	4
3	54	Private	140359	7th-8th	4	Divorced	Machine-op-inspct	Unmarried	White	Female	0	3
4	41	Private	264663	Some-college	10	Separated	Prof-specialty	Own-child	White	Female	0	3
...	...	...	...	...	...	...	...	...	...	...	...	
32556	22	Private	310152	Some-college	10	Never-married	Protective-serv	Not-in-family	White	Male	0	
32557	27	Private	257302	Assoc-acdm	12	Married-civ-spouse	Tech-support	Wife	White	Female	0	
32558	40	Private	154374	HS-grad	9	Married-civ-spouse	Machine-op-inspct	Husband	White	Male	0	
32559	58	Private	151910	HS-grad	9	Widowed	Adm-clerical	Unmarried	White	Female	0	
32560	22	Private	201490	HS-grad	9	Never-married	Adm-clerical	Own-child	White	Male	0	

32561 rows × 15 columns

Next steps: Generate code with `df` View recommended plots

```
df1 = df[['age','fnlwgt','education.num','capital.gain','capital.loss','hours.per.week']].copy()
df1
```

	age	fnlwgt	education.num	capital.gain	capital.loss	hours.per.week
0	90	77053	9	0	4356	40
1	82	132870	9	0	4356	18
2	66	186061	10	0	4356	40
3	54	140359	4	0	3900	40
4	41	264663	10	0	3900	40
...	...	...	...	...	...	...
32556	22	310152	10	0	0	40
32557	27	257302	12	0	0	38
32558	40	154374	9	0	0	40
32559	58	151910	9	0	0	40
32560	22	201490	9	0	0	20

32561 rows × 6 columns

Next steps: Generate code with `df1` View recommended plots

```
df2 = df1
for column in df2.columns:
  df2[column] = (df2[column]-df2[column].min())/(df2[column].max()-df2[column].min())
df2
```

	age	fnlwgt	education.num	capital.gain	capital.loss	hours.per.week
0	1.000000	0.043987	0.533333	0.0	1.000000	0.397959
1	0.890411	0.081896	0.533333	0.0	1.000000	0.173469
2	0.671233	0.118021	0.600000	0.0	1.000000	0.397959
3	0.506849	0.086982	0.200000	0.0	0.895317	0.397959
4	0.328767	0.171404	0.600000	0.0	0.895317	0.397959
...	...	...	...	...	...	...
32556	0.068493	0.202298	0.600000	0.0	0.000000	0.397959
32557	0.136986	0.166404	0.733333	0.0	0.000000	0.377551
32558	0.315068	0.096500	0.533333	0.0	0.000000	0.397959

Next steps: Generate code with df1 View recommended plots

```
df3 = df1
for column in df3.columns:
  df3[column] = (df[column]-df[column].mean())/df[column].std()
df3
```

	age	fnlwgt	education.num	capital.gain	capital.loss	hours.per.week
0	3.769554	-1.067981	-0.420053	-0.145918	10.593344	-0.035429
1	3.183063	-0.539160	-0.420053	-0.145918	10.593344	-1.817176
2	2.010079	-0.035219	-0.031360	-0.145918	10.593344	-0.035429
3	1.130342	-0.468208	-2.363521	-0.145918	9.461719	-0.035429
4	0.177293	0.709471	-0.031360	-0.145918	9.461719	-0.035429
...	...	...	...	...	...	...
32556	-1.215625	1.140442	-0.031360	-0.145918	-0.216656	-0.035429
32557	-0.849067	0.639731	0.746028	-0.145918	-0.216656	-0.197406
32558	0.103982	-0.335428	-0.420053	-0.145918	-0.216656	-0.035429
32559	1.423588	-0.358772	-0.420053	-0.145918	-0.216656	-0.035429
32560	-1.215625	0.110958	-0.420053	-0.145918	-0.216656	-1.655199

32561 rows × 6 columns

Next steps: Generate code with df1 View recommended plots

Start coding or generate with AI.