In [35]:
```
from completejourney_py import get_data

cj_data = get_data()

import pandas as pd
transactions = cj_data['transactions']
products = cj_data['products']
demographics = cj_data['demographics']
transactions.columns.intersection(demographics.columns)
```

Out[35]:
```
Index(['household_id'], dtype='object')
```

In [41]:
```
df1 = transactions.merge(demographics, how='left', on='household_id')
```

In [43]:
```
df1.columns.intersection(products.columns)
```

Out[43]:
```
Index(['product_id'], dtype='object')
```

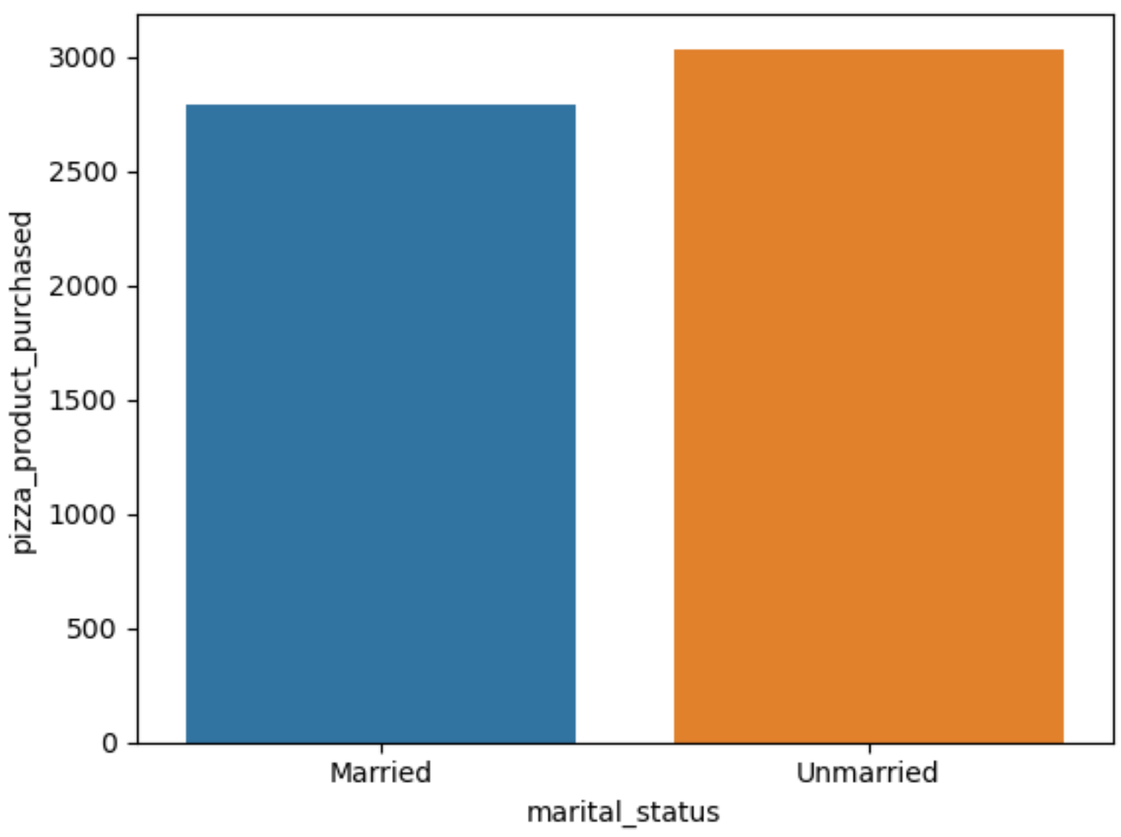
In [96]:
```
pizza_filter = products['product_type'].str.contains('pizza', case=False, na=False)
df2 = (products[pizza_filter].merge(df1 , how='inner', on='product_id')
       .groupby(['marital_status'], as_index=False)
       .agg({'product_type': 'count', 'sales_value': 'sum'})
      )
df2 = df2.rename(columns={'product_type': 'pizza_product_purchased'})
df2
```

Out[96]:

	marital_status	pizza_product_purchased	sales_value
0	Married	2788	10711.37
1	Unmarried	3037	10917.73

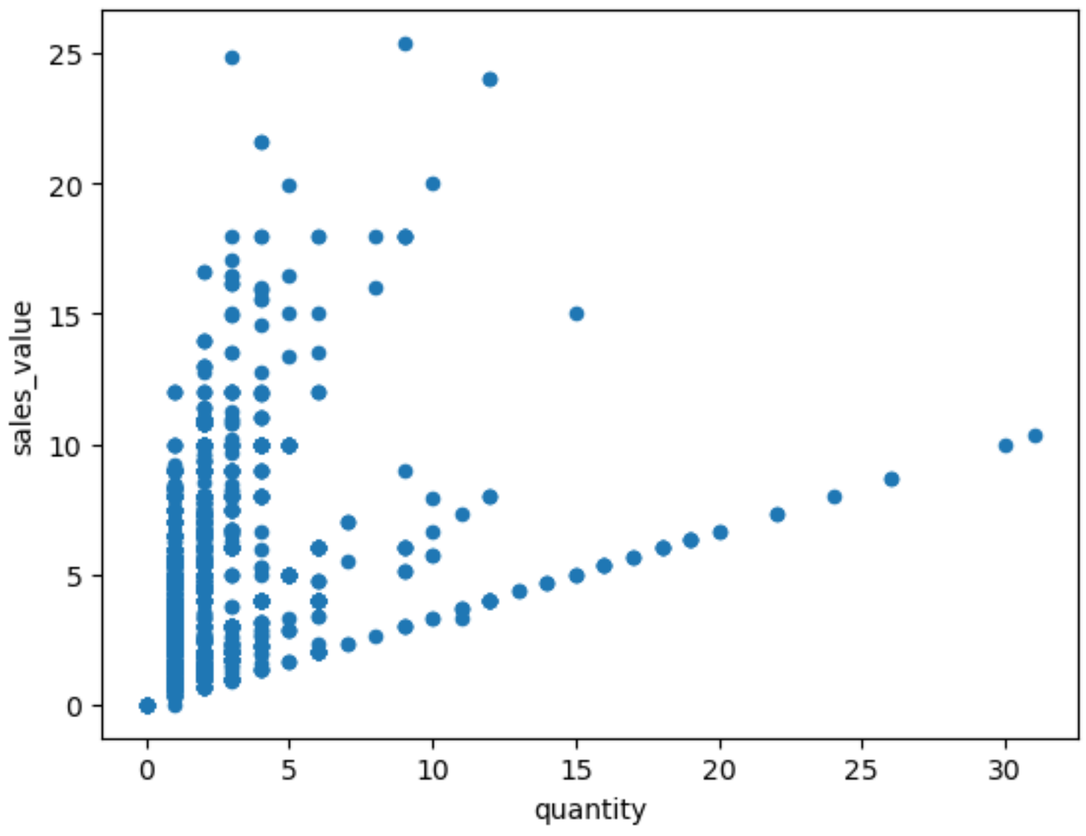
In [94]:
```
import seaborn as sns
import matplotlib.pyplot as plt
sns.barplot(x='marital_status', y='pizza_product_purchased', data=df2)
```

Out[94]:
```
<Axes: xlabel='marital_status', ylabel='pizza_product_purchased'>
```



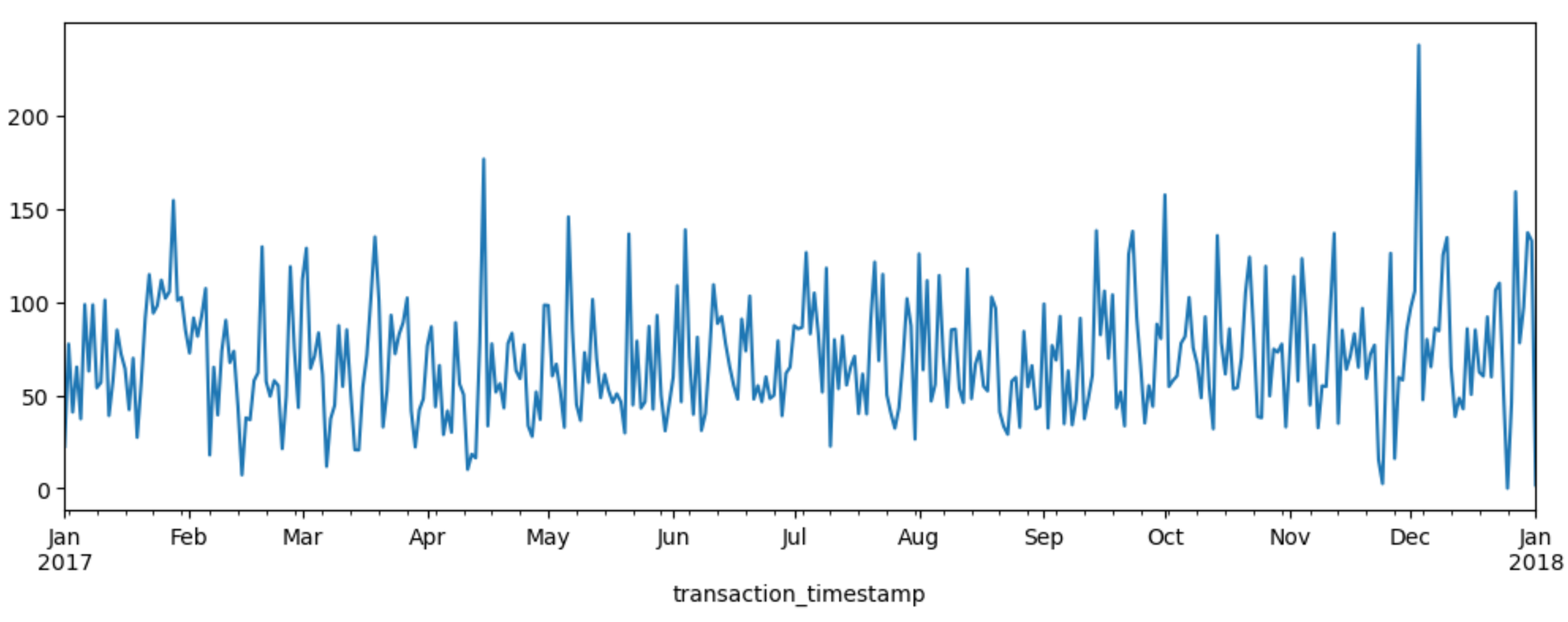
In [106…
```
df3 = (products[pizza_filter]
       .merge(transactions , how='inner', on='product_id')
       .merge(demographics, how='inner', on='household_id')
      )
df3.plot.scatter(x='quantity', y='sales_value')
```

Out[106…
```
<Axes: xlabel='quantity', ylabel='sales_value'>
```



In [112…
```
df4 = df3.set_index('transaction_timestamp')['sales_value']
sale_resample = df4.resample('D').sum()
sale_resample.plot(kind='line', figsize=(12, 4))
```

Out[112…
```
<Axes: xlabel='transaction_timestamp'>
```



In [ ]:
```

```