Conditional probability using pandas

Special thanks to Roba, Data Analysis Session Lead for referring the dataset

Dataset

Telco Customer Churn | Kaggle
https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Read through the dataset description and then answer the following questions. You can either download the dataset and work locally or you can work in a Kaggle notebook

Tasks

1- check for missing values
2- check for duplicates
3- check for categorical variables that have unusual categories
4- Choose some factors that may affect customer churn and test the P(Churn|factor)
5- can you use value_counts to get proportions? If yes, can we use it here?
6- Can pandas crosstabs do all steps in one line? Implement this **Hint:** check the function's arguments [use the shift + tab option]
7- check `Minimally sufficient pandas` for a comparison between crosstabs vs pivot table vs groupby

Minimally Sufficient Pandas - A guide to use pandas effectively
https://www.dunderdata.com/blog/minimally-sufficient-pandas

Minimally-Sufficient-Pandas/9. Similarity between groupby, pivot_table, crosstab.ipynb at master · tdpetrou/Minimally-Sufficient-Pandas
https://github.com/tdpetrou/Minimally-Sufficient-Pandas/blob/master/9.%20Similarity%20between%20groupby%2C%20pivot_table%2C%20crosstab.ipynb

8- present your work