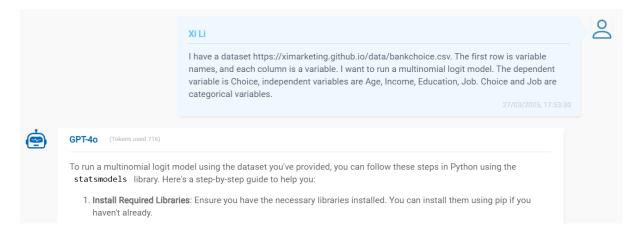
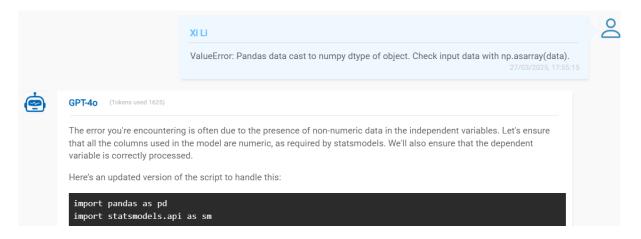
I have a dataset https://ximarketing.github.io/data/bankchoice.csv. The first row is variable names, and each column is a variable. I want to run a multinomial logit model. The dependent variable is Choice, independent variables are Age, Income, Education, Job. Choice and Job are categorical variables.



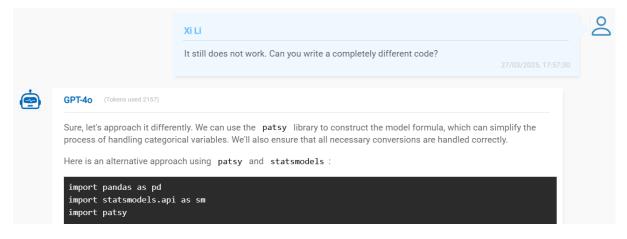
## KeyError: 'Choice\_1'

		Xi Li			0
		KeyError: 'Choice_1'			
ė	GPT-40 (Tokens used 1116)				
	It looks like there's an issue with how the dependent variable 'Choice' is being referenced in the script. The dataset most likely has the 'Choice' column as a categorical variable, and it needs to be converted correctly.				
	Here's an improved version of the script that should handle the conversion of the 'Choice' variable correctly:				
	<pre>import pandas as pd import statsmodels.api as sm from statsmodels.tools import add_constant</pre>				

ValueError: Pandas data cast to numpy dtype of object. Check input data with np.asarray(data).



It still does not work. Can you write a completely different code?



It works this time!