Forage Data Analytics with GrenAI.

Taski: EDA (Exploratory Data Analysis)
Techniques to handle missing values leverage synthetic data generation when data is limited.

Predict (redileard deliquency.

Need of Company: More structured, data driver approach current method: historical trends & broad segmentation maly AI diver insights kmodel should be made

My work:

1) Review doitaset = completeners Key attributes

missing incorrect values identify

②Use GenAI Tools: generale insights ensure confidentiality

3 Sunnaige findings: pattens

Key visk indicators report

*EDA: identity patterns, trends, inconsistencies, missing values to ensure data quality and reliability.

* Data integrity: correct data without missing values, duplicates, inconsistencies.

* steps in EDA:

@Understanding dataset_ Key variables categorical

inconsistent values numerical data points

Use Gren AI tool

prond: malyze this dotaset & provide a summary of key columns, including common patterns and missing values.

@ Identify missing values koutliers: techniques 1) Statistical imputation: Replace missing values with moon, median or regression d. based imputation. 2 Understanding missingness patterns: data is Missing completly at random (MCAR) Missing at random (MAR) Musing not at a random (MNAR) Doll Removing Irrelavant data: ensure no bias. Gent I prompt - Identity inissing values in this dataset and recommend the best infutation strategy based on industry best practices. 3 Urderstanding relationships botuseen wriables: GenAI prompt :- Analyse the correlation bow customer income and delinquency visk summarizing Key findings in simple terms S theek too displicable, in 4 Detecting patterns Krisk factors Gents prompt: Avalyze trunds un late payments and identify the top 3 risk factors associated with delinquency Use Neam = cit dataies well balanced without externe values Use Median= it data contains outliers that effects average Use Node = et single value is repeated many times. Applying A GenAI generated imputation without understanding data context can introduce gignificant bias. Systectic dated Otherway small dateste Orithing in missing phillips sew of merico Slebel In priting @

(4) Evening many conflict

* Techniques for hardling missing values of your productions
* Techniques for hardling missing
6 1 1 1 000 000 000 000 000 000 000 000
mandy a company dataster auce
Demolon evers
Ockwed data collection
Okandon evrous Okandon evrous Oskewed data cellection Oskewed data cellection Oskewed behavious
Consult material to too they grisself
How to harrolle: wissing data miles
1 Deleting vision
@ Imputant Mean, Median, Moord filling - Uses
polarie mistra data trende to extinate mising
① Deleting missing. ② Imputation — Mean, Median, Mode — Forward or backward filling - uses — Forward or backward to estimate musing existing data tends to estimate musing entries.
3 AI assisted Imputation.
1 doket patterns & prouder sugarion.
General pursual to an experience of etatistical empressed in the contractions
3 AI assisted Imputation. Loket patterns & provides suggestions of statistical imputation. Suggest synthetic data Suggest synthetic data
> check to duplicates, inconsistent formatting & logical ersors.
=> check for desplicales, inconsiste Here & sensitive
Check to duplicates, inconsistent formatting be regical consistent synthetic data generation is limited or sincomplete, synthetic data generation is done to till gaps. Synthetic data generation is done to till gaps. Synthetic data generation is done to till gaps. Synthetic data generated data Synthetic da
une to fill gaps.
conthetic data generation as me
merated data The merated data
attrically go.
outsticially generated data the constant for real data. This will provide security privacy for real data.
this will provide sources of Tookmiques -
statistical simulation letters
-> Hadition to Monte coulo a simulation biling Tomes to
this will provide servity privacy for real dotted this will provide simulation Techniques: Traditional Statistical Simulations (2) Rootstropping (3) probabilistic modeling When to use synthetic data; (1) Enhancing small datasets (2) Filling in missing values (3) Tothing of Models
(3) probabilitie mounty
When to use synthesis activing in missing values
@ Testing AI Models
(9) Ensuing Pulsacy compliance

GenAII Exaufile proufd? Generate synthetic payment history data for customers with missing vecolds while ensuring that distributions align with historical patterns observed in the dataset (e.g., standard deviation, typical payment behaviors).

* Make sure synthetic data doesn't introduce bias.

** Use guadratic free Tool for getting insightsk

analysis.

pisson

endi

48.

.

riques

آ سه

lavce