                                                            Data Privacy Assignment

**Fin Assist: Privacy-first digital financial advisor platform**

This dataset means that it basically appears to be digital financial advisory user feedback dataset that investigates into the demographic, financial, Investment and advisory and finally goals and actions  feedback and trust.

Attributes : There are multiple attributes or columns in this dataset.

When we split the columns on the basis Demographic and Personal data

Name : gives  user’s name Attributes Type String (Text) Atrributes : names

Age: gives the user’s age Attribute type is Integer

Location : Gives the user’s Location or can be an Address  Attribute type is String

Gender : which provides the gender of the user Attribute  type string

Here we work on the Financial Profile Profile

Income : the net income of customer

asset & debt  : analyzing the level of assets and debt the user has and help the advisor to personalize recommendation

Monthly Expanses : Average monthly expenses of the user Reflects spending behavior type is Numerical(Float)

Type of Advice Sought : Shows the financial advice areas users are interested in type is string(categorical) attributes :        Tax Planning ,Retirement Planning

User Financial Goal: May reflects specific financial priorities indirect indentifier type: string(Text) attributes: Tax savings , Debt Reduction

Action Based on Advice : Tracks the users responce to the advice given. Type:String(categorical) attributes:  Implemented , Partially Implemented

Feedback on Advice : Reflects satisfaction levels could contain free form text exposing personal information Type: String(Text) attributes: Very unsatisfied , neutral, satisfied

App Usability Feedback :  Users preception of the apps performance and ease of use type: String(Categorical) Attributes : Unsatisfied , Satisfied

Trust Level with Digital Advisor : Indicates user trust and confidence in the digital advisor , revealing personal preferences Type :  String(Categorical) attributes: Low Moderate High

Investment Prefernece : Inticate user interests in investment options provides insight into fincial behavior type string  attributes : Fixed Deposits Gold Real Estate.

Direct Identififiers the name column is a direct identifier.

Indirect Identifiers Age, Location , Gender, Income can combine and give a data that could lead to the identification of specific users.

Overall this dataset requires careful handling due to the mix of direct identifiers, indirect

Sensitivity financial data Proper anonymization, encryption will keep the privacy.

A diagram of a data storage system

Description automatically generated

Threat category #1 (Likability)

Hotspot: User information during registration, Users financial behavior analysis, Transaction data from Implemented actions

Specific threat: Name and location data can be linked to other database exposes the user identity Name and Location fields needs to be cross referred with social media to establish connections like Monthly Expenses linked with Income can tell users spending habits and can be used for financial scams.   Action Based on Advice can be traced back to individuals, connecting them to specific transition.

Threat category #2 (Identifiability)

Hotspot : Data storage of personal details , Users investment preferences, Custom financial goal creation

Specific Threats: Identifiable fields like Name and Location are stored without anonymization. Fields like Name Age, and Gender are personal if not properly used they can be misused.  Investment Preference and Type of Advice Sought could reveal specific financial interests. User Financial Goal can reflect personal ambitions.

Threat category #3 (Non- repudiation)

Hotspot : Users action tracking within the advisor, Tracking users interactions with the financial advisor, Users feedback entries, Response to financial guidance, Incomplete tracking of investment actions

Specific Threats: Lack of logs for advice can lead to disputes on users actions. The absence of secure and timestamped logs of advice given can lead to the deniability on the services provided. Lack of verification on Feedback on Advice allows user to deny the submission of feedback. Lack of digital signature for advice responces can make it difficult. Inconsistency in Record keeping for Investment Preference can lead to dispute regarding provided recomandation

Threat category #4 (Detectability)

Hotspot : Financial advice search activities, Analysis patterns in Investment Preference Finacial Goal, Monitoring trust levels with the digital advisor, Patterns in app usage feedback, User demographics in trageted advice, Observing changes in financial Priorities

Specific Threats: Search and selection patterns can be tracked if not anaonymzed. If the data is not anonymized user behavior paterns could lead to be dected and exploit by third parties. Changes in Trust Level with Digtal Advisor could indicate dissatisfaction allows malicious attack on target vulnerable users. App Usability Feedback could use exploitable weekness in the application. Combination of Age and Gender with financial goals allows for profilling and trageting the users data. Monitoring shits in User Finacial Goal might make personal changes exposes the pivate data.

Threat category #5 (Disclosure of information)

Hotspot : Income and assets data, comments in feedback on Advice, Debt data during profile setup

Specific Threats : Sensitive finacial data like Income and Assets can be expossed if not encrypted, Free form text in feedback could unitentionally reveal sensitive personal or financial information if share too much information. Exposure of Dept field could lead to social engineering attacks targetting users.

Threat category #6 (Unawareness)

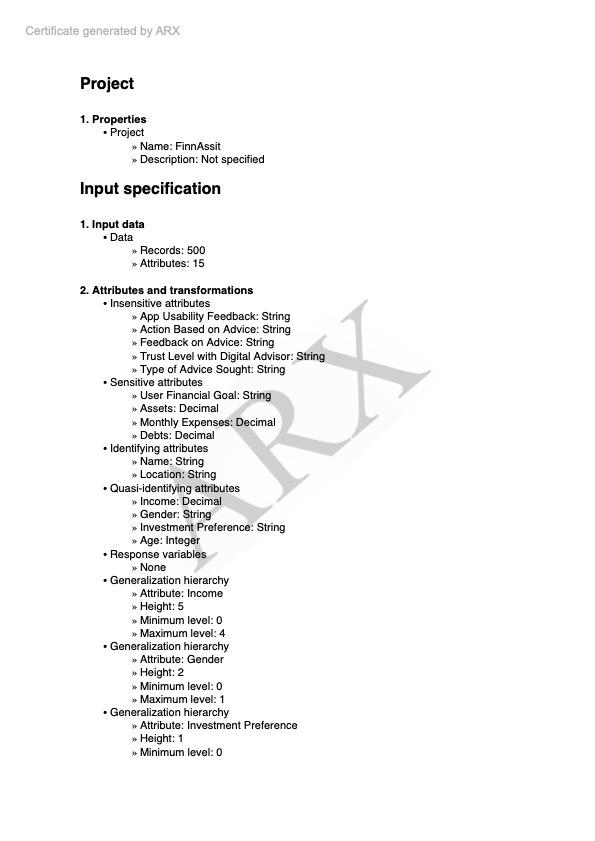
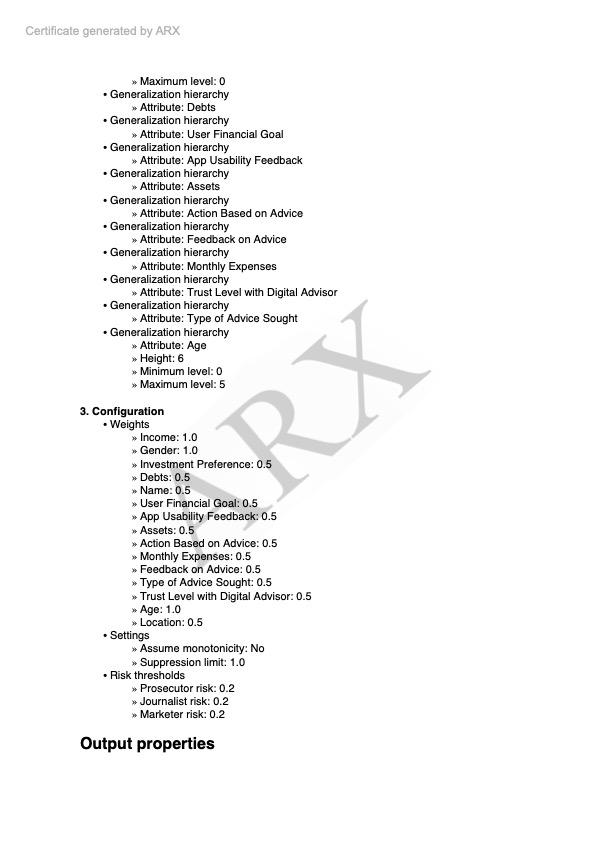
Hotspot : Account Registeration with weak credentials , Sharing details financial information, Misunderstanding data privacy settings

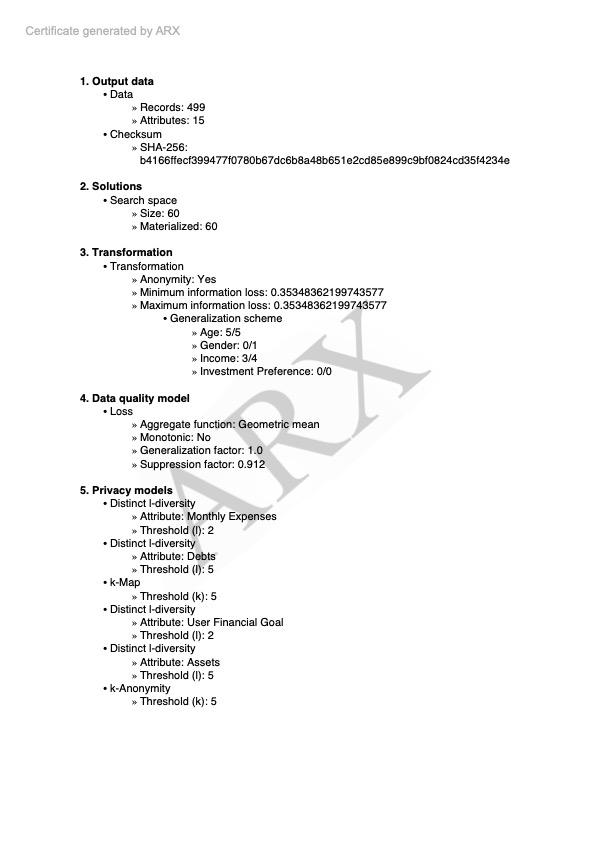
Specific Threats : Users may use simple passwords and making account vulnerable to brute-force or stuffing attacks. Users might provide excessive information in free spaces like Feedback on Advice without realizing risks of exposure. Users might not configure privacy options properly, unintentionally allowing third parties to access Investment Preference and Financial Goal data.

Threat category  #7 (Non-compliance)

Hotspot : Failure to follow data encryption guidelines , Ignoring multi-factor authentication requirements , Not implementing secure password polices

Specific Threats : Financial data like Income and Assets mmay not be encrypted as per security standards, leading to potential data breaches. Users may not be requried to set up multi-factor authentication(2FA) increasing the risk of unauthorized account access. Users could register with weak passwords rules, exposing accounts to brute-force attacks.



Here is an explanation of the decisions taken based on the ARX certificate analysis:

Input Specifications:

Data: The dataset consists of 500 records with 15 attributes.

Attributes and Transformations:

Insensitive Attributes: App Usability Feedback, Action Based on Advice, Feedback on Advice, Trust Level with Digital Advisor, Type of Advice Sought.

Sensitive Attributes: User Financial Goal, Assets, Monthly Expenses, Debts.

Identifying Attributes: Name and Location (treated as highly identifying for privacy protection).

Quasi-Identifying Attributes: Income, Gender, Investment Preference, Age.

Response Variables: None specified.

Generalization Hierarchy: Attributes such as Age have 6 levels of generalization, while Gender and Income have more limited hierarchies.

Configurations:

Weights: Each attribute is assigned a weight to reflect its importance and sensitivity (e.g., Income, Gender, and Age carry higher weights of 1.0, while other attributes like Trust Level with Digital Advisor are weighted at 0.5).

Settings: Monotonicity assumption is set to 'No' with a suppression limit of 1.0.

Risk Thresholds: Set for different attacker models, including Prosecutor, Journalist, and Marketer risks, all capped at 0.2.

Output Properties:

Output Data: 499 records and 15 attributes (indicating minimal data suppression).

Solutions: Search space size of 60 solutions with all materialized.

Transformations: Anonymity was achieved with minimum and maximum information loss set at 0.353. Generalization schemes for attributes like Age (maximum level 5/5), Income (3/4), and Gender (0/1) were applied.

Data Quality Models: Used a geometric mean aggregate function for assessing data loss, with a generalization factor of 1.0 and a suppression factor of 0.912.

Privacy Models:

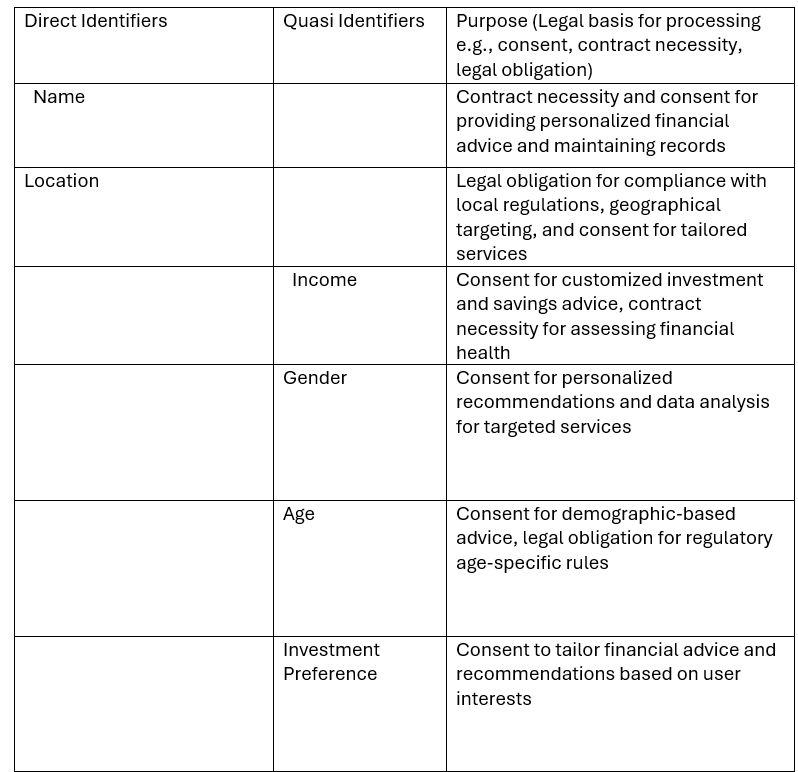
k-Anonymity: Achieved with a threshold of k=5.

Distinct l-Diversity: Applied to attributes such as Monthly Expenses (l=2), Debts (l=5), User Financial Goal (l=2), and Assets (l=5).

k-Map: Threshold set at k=5.

These decisions aimed to balance data utility with privacy through comprehensive anonymization and transformation strategies to meet privacy standards while minimizing information loss.

**spreadsheet of data inventory**



**Data Processing Register**

