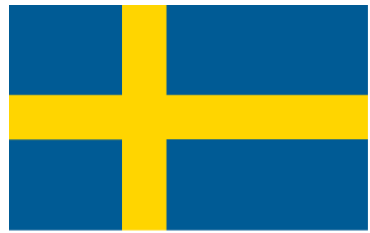**ФОНД ГАРАНТУВАННЯ ВКЛАДІВ**

Summary report

Study on pricing factors and price prediction modelling for sales of the assets of bankrupted banks based on historical sales conducted by the DGF on Prozorro.sale in Ukraine

January 2024

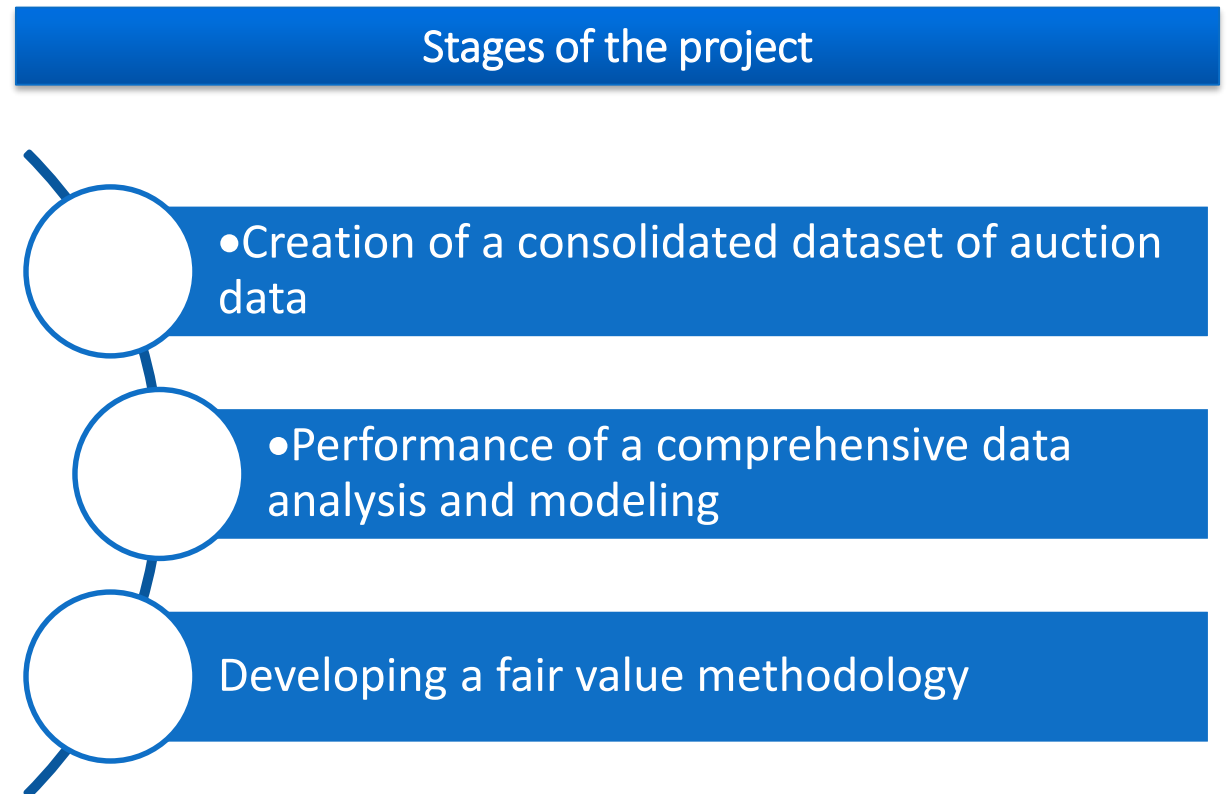# Project Sponsors

Content

Project Overview

Project Overview

The cornerstone of this project is to conduct a detailed study of historical sales transactions done by the DGF during the period from 2017 to February 2022 and prepare a fair price methodology to be used during the resolution of insolvent bank for estimation of the future incoming cash flows from divesting assets in liquidation procedure.

The work, performed by one of the Big 4 companies, is based on best practices and methods adapted to the Ukrainian market. The first step of the project is creation of a consolidated dataset of asset sales conducted by DGF. On the next steps collected data is categorized and analyzed to create a basis for further project stages. As a result of this study, predictive models for calculation of assets fair value are developed to be used by DGF during performance of its function.

Stages of the project

- Creation of a consolidated dataset of auction data

- Performance of a comprehensive data analysis and modeling

- Developing a fair value methodology

Project Overview

Main project components →

Data collection, consolidation, and analysis

- Prepare (clean and consolidate) a large pool of the DGF's and Prozorro.Sale's data related to asset sales conducted by the DGF. Perform data segmentation based on the type of assets sold.

- Identify variables that impact the final sale price, successful auction timing, or factors that may cause unsuccessful auctions.

- Identify sale price trends, including macrotrends, for the period, for different asset classes, for lot types (individual/bulk).

- Perform a sensitivity analysis of the sale price regarding different parameters.

Development of the models

Develop statistical models for each type of assets (if possible) based on available historical data and data analysis results. The main assets are:
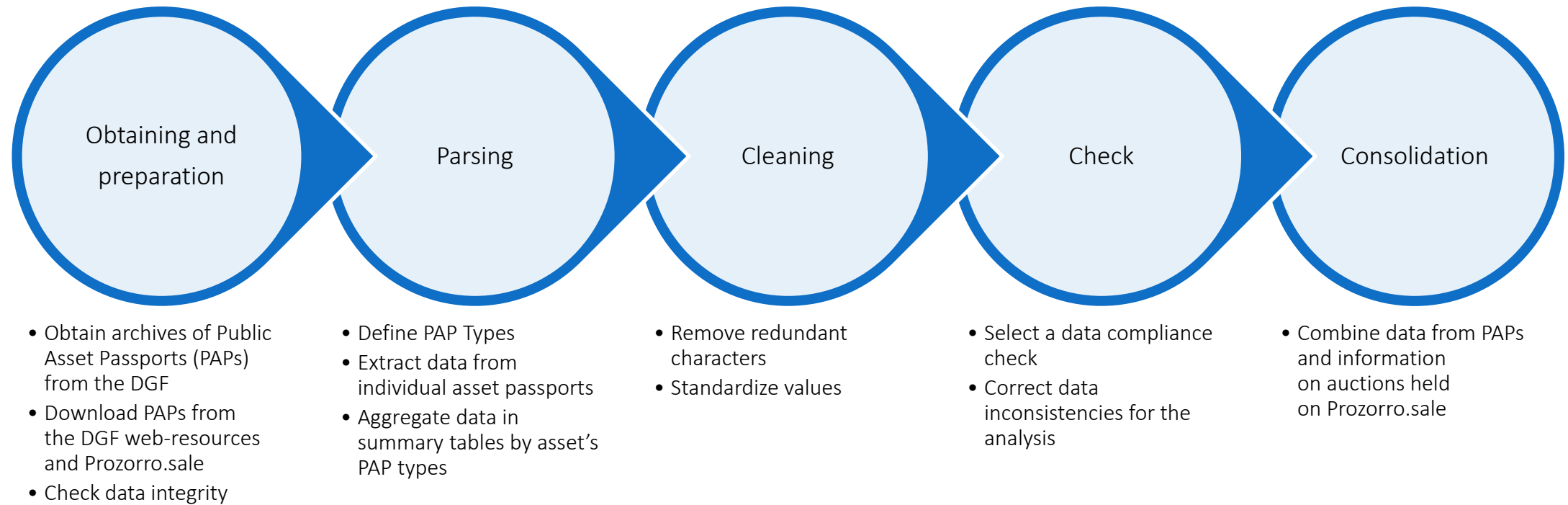
- loans

- vehicles

- fixed assets

- real estate

- securities

Fair value methodology
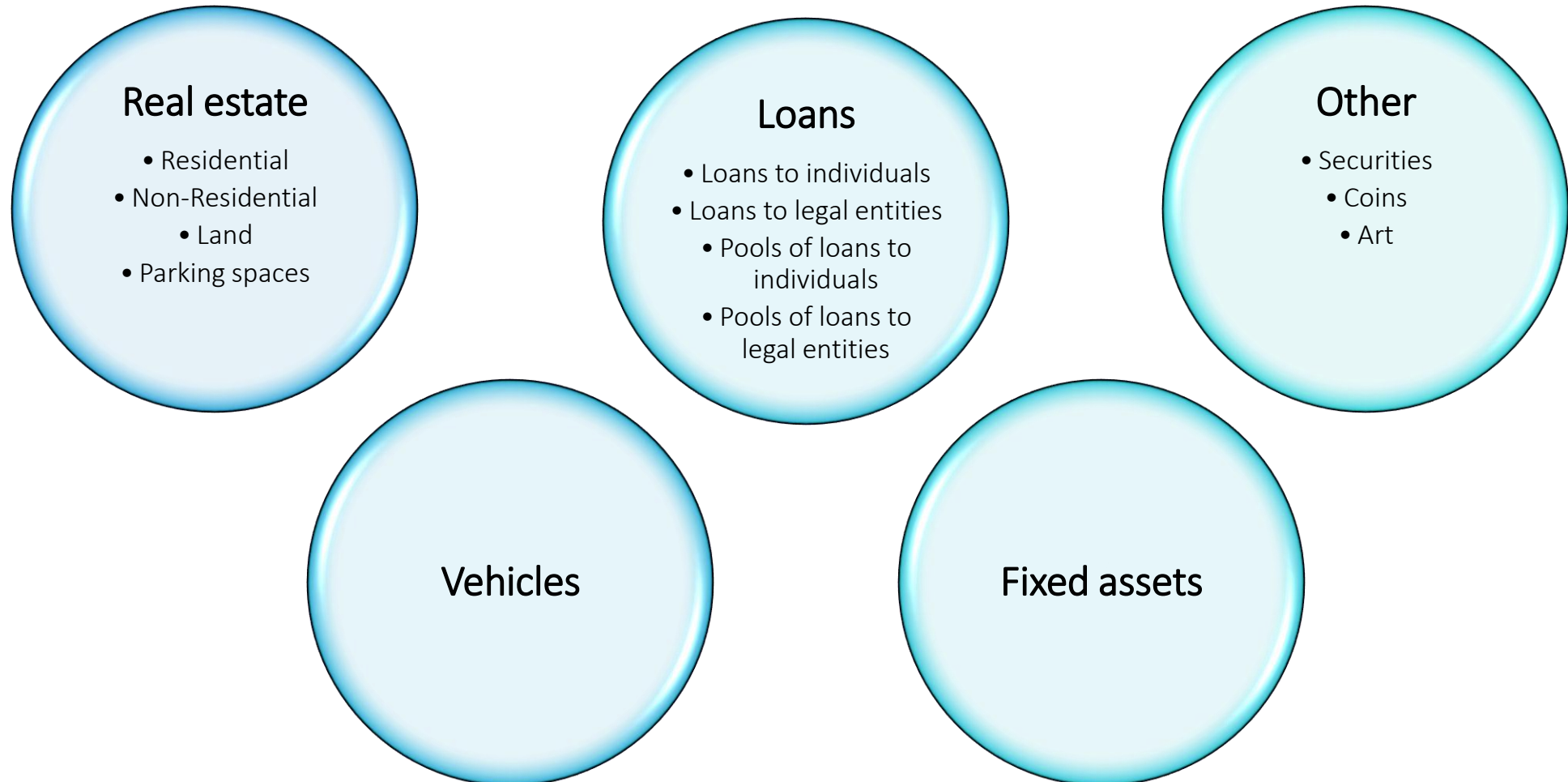
- Prepare a methodology for estimating future cash flows from sales when divesting assets.

- Provide a list of available valuation methods with indication of their applicability for a certain asset class (i.e., loans, commercial real estate, land, etc.).

- Incorporate relevant findings of the asset sale study conducted at the first stage of the project.

Creation of a Consolidated Dataset

Creation of a Consolidated Dataset

Obtaining and preparation

- Obtain archives of Public Asset Passports (PAPs) from the DGF
- Download PAPs from the DGF web-resources and Prozorro.sale
- Check data integrity

Parsing

- Define PAP Types
- Extract data from individual asset passports
- Aggregate data in summary tables by asset's PAP types

Cleaning

- Remove redundant characters
- Standardize values

Check

- Select a data compliance check
- Correct data inconsistencies for the analysis

Consolidation

- Combine data from PAPs and information on auctions held on Prozorro.sale

Creation of a Consolidated Dataset – Defined Asset Types

Real estate
- Residential
- Non-Residential
- Land
- Parking spaces

Loans
- Loans to individuals
- Loans to legal entities
- Pools of loans to individuals
- Pools of loans to legal entities

Other
- Securities
- Coins
- Art

Vehicles

Fixed assets

Creation of a Consolidated Dataset - Parsing

Original PAP

ПАСПОРТ АКТИВУ Транспортні засоби та спеціалізована техніка	ФГВФО
Назва банку	АТ "Брокбізнесбанк"
1. Характеристика майна	
1.1. Назва активу: транспортні засоби чи спецтехніка	Транспортні засоби
1.2. Вид транспортного засобу (легковий автомобіль, вантажний автомобіль, автобус, літак, судно, потяг, вагон, причіп, платформа, трейлер, інше) чи вид спецтехніки (бульдозер, трактор, комбайн, автокран, погрузчик, ескаватор, тощо)	Автомобіль д.н. BB0011BB
1.3. Марка та модель транспортного засобу/спецтехніки	Toyota Camry 2.0
1.4. Рік випуску	2006
1.5. Об'єм двигуна	3456
1.6. Пробіг (км) чи наработка (моточаси)	332396
1.7. Колір	жовтий
1.8. Номер кузова/шассі/VIN-код	4T1BK46K47U506139
1.9. Фізична наявність (наявний, розшук, арешт тощо)	наявний
1.10. Фізичний стан (відмінний, добрий, задовільний, незадовільний)	задовільний
опис пошкоджень	0
1.11. Місце реєстрації за свідоцтвом про реєстрацію	Луганськ

Consolidated table

Вид транспортного засобу	Марка та модель транспортного зас	Рік випуску	Об'єм двигуна	Пробіг (км) чи наработка	Колір	Номер кузова/шассі/VIN-код
Автомобіль АА6367ВЕ,	ГАЗ 2217	2005	2300	236720	білий	інформація відсутня
Автомобіль д.н. АА2647ІС	Volkswagen transporter інкасаційний	2008	2461	352307	сірий	WV1ZZZ7HZ8H159460
Автомобіль д.н. BB0011BB	Toyota Camry	2006	3456	332396	жовтий	4T1BK46K47U506139
Автомобільд.н. АА7393КР	Volkswagen T4 інкасаційний	2003	2461	658820	сірий	WV1ZZZ70Z3H139917
Легковий Сєдан	TOYOTA CAMRY	2007	2362	173000	сірий	JTNBE40K403091397
Легковий Седан-В	DAEWOO LANOS	2012	1498	162685	сірий	Y6DTF69Y0D0318358
Легковий Седан-В	DAEWOO LANOS	2012	1498	194387	сірий	Y6DTF69Y0C0307835
Легковий-В Седан	DAEWOO LANOS TF69Y	2005	1498	308598	сірий	SUPTF69YD5W239550
Легковий Седан-В	SKODA SUPERB ELEGANCE	2007	1781	147001	чорний	TMBBL23U08B300065

Details of assets were extracted using the following technologies:
- Robotic Process Automation (RPA)
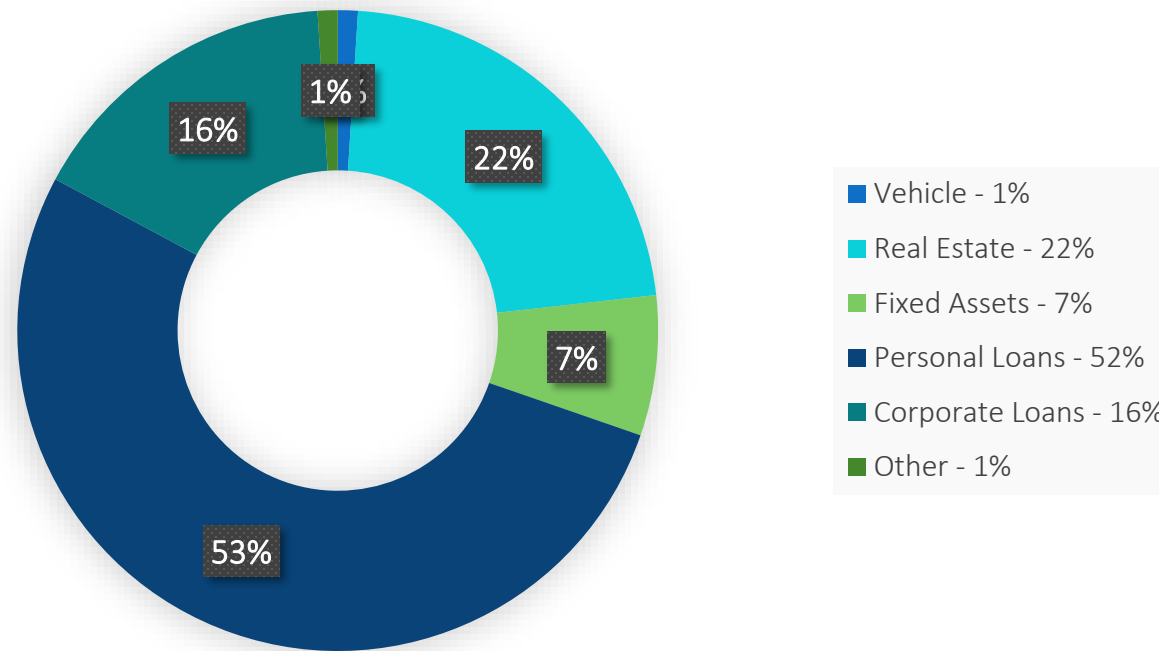- The parsing technique implemented in Python

Information is extracted from PAPs using developed identifiers and is placed in the corresponding column of the tables, as shown in the illustration.

The above example demonstrates the extraction of the model of the vehicle and the VIN for PAP template of vehicles..

Creation of a Consolidated Dataset – Asset Types Distribution

The total number of conducted auctions for the period from 2017 to 2022 was **221,952**, for which **88%** of Public Asset Passports were obtained from public available sources. The majority of data relates to Loans and Real Estate, the rest of asset types account for less than 10% of the overall number of auctions.

Asset Types Distribution

Legend:
- Vehicle - 1%
- Real Estate - 22%
- Fixed Assets - 7%
- Personal Loans - 52%
- Corporate Loans - 16%
- Other - 1%

Chart values shown: 1%, 22%, 7%, 53%, 16%, 1%

Data Analysis and Model Building

Price Prediction Model – Data used for modeling

Since the templates of the public asset passports and the datasets they contain changed during the analysis period and were not identical in content, the datasets underwent stages of cleaning and validation. Although such data can and should be stored for future reference, it is not suitable for the purposes of this study and was excluded from the dataset used for sales analysis and model.

Observations that contain complete and reliable information about assets were used for further analysis and price prediction modeling. The total number of observations (within all asset classes) in the "Sold assets with cleaned data" sample is 5,827, which is 52% of the total number of assets sold.

Total number of auctions
221,952

Unique lot numbers
83,278

Assets sold
11,100

Sold assets with cleaned data
5,827
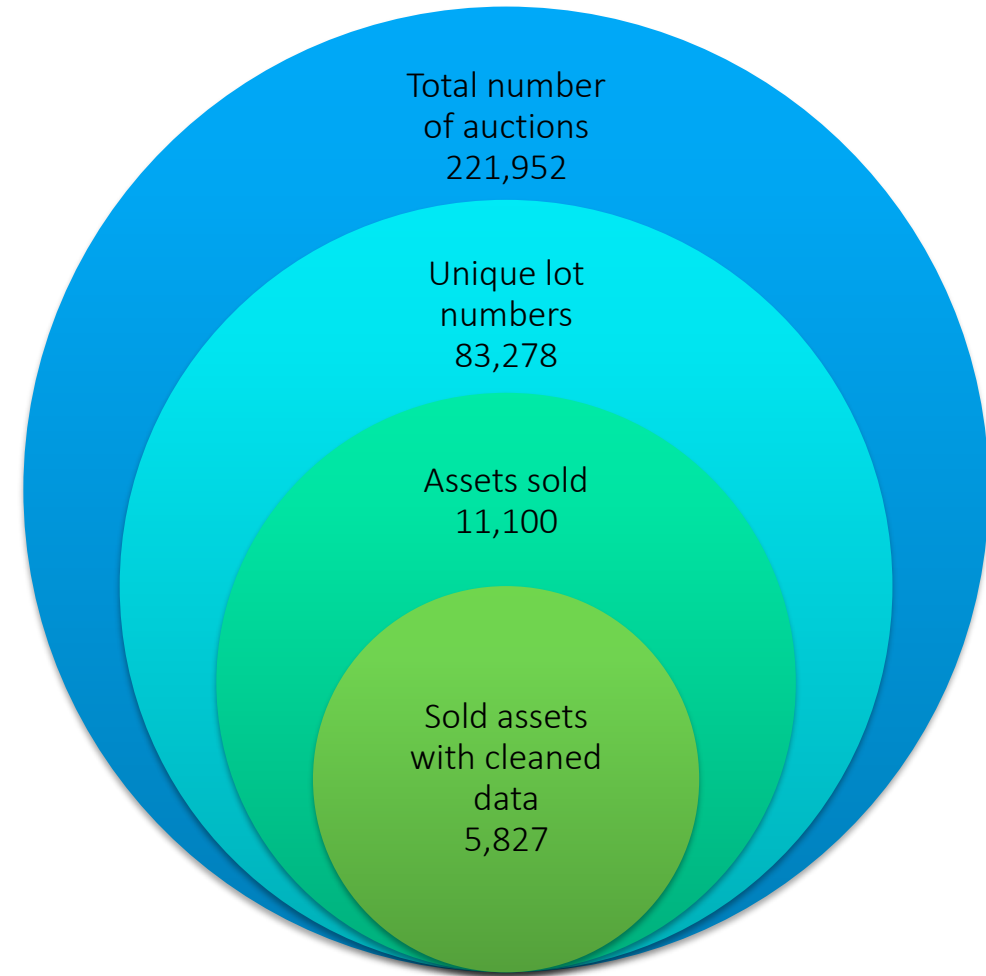
Price Prediction Model[1] – Vehicles

	Factor	K_i
1	Engine Volume	0.167795
2	Physical Availability – Seized/Traced	-0.14947
3	Age (square root)	-0.55096
4.a	Segment – Budget	-0.21675
4.b	Segment – Middle	-0.0855
4.c	Segment – Upper middle	-0.00975
4.e	Segment – Luxury	0.311997
5.a	Brand Region – USA	0.025173
5.b	Brand Region – Other countries	-0.59112
5.c	Brand Region – Europe	0.250328
5.e	Brand region – Asia (South Korea and Japan)	0.315615
6	Mileage	-0.04891
7	Physical condition – Unsatisfactory	-0.28722
8	Type of transport – Cars	0.121301
	Intercept (I)	9.919782

$$P = \sum_i K_i \cdot V_i + I$$

A statistical model has been developed to predict the selling price of vehicles. Based on the results of the analysis of the dataset, a linear model with modification—*Ridge Regression (α=0.01)*—was chosen.

This model calculates the expected selling price—P—based on the 8 factors shown in the table. *Engine Volume, Age (square root),* and *Mileage* are quantitative factors, which means that they have numerical values V_i that affect (with respective K_i ratios) the predicted sales price. *Physical Availability – Seized/Traced, Segment (Budget, Middle, Upper middle, Luxury), Brand Region (USA, Europe, Asia (South Korea and Japan), Other countries (all other countries not included in the regions above), Physical condition – Unsatisfactory, Type of transport – Cars* are qualitative factors and take "1" for "Yes" or "0" for "No", in which case and based on K_i ratio, they affect or do not affect the estimated price of the asset.

MAPE	R2
0.21	0.76

Price Prediction Model[1] – Fixed Assets (FA)

$$P = Quoted\ price \cdot Discount$$

P – prognosed income for the FA pool,

$Quoted\ price$ – the maximum value between the Book Value and the Estimated Value of the FA pool (UAH),

$Discount$ – the value of the discount (statistically calculated) depending on the type of fixed assets.

Fixed Assets Type	Discount
Banking equipment	37%
Furniture	54%
Electronics	50%
Other	53%

- Generally, fixed assets are sold in pools. In most cases, fixed assets are sold in pools or in combination with other assets (for example, real estate, vehicles). The developed model covers only pools for the sale of fixed assets.

- Among asset characteristics specified in the passports, one has the greatest impact on the cost of fixed assets—the fixed asset type.

- The limited number of asset characteristics in passports allows building the simplest model, which is based on sample metrics and takes into account just one factor, i.e. the fixe asset type.
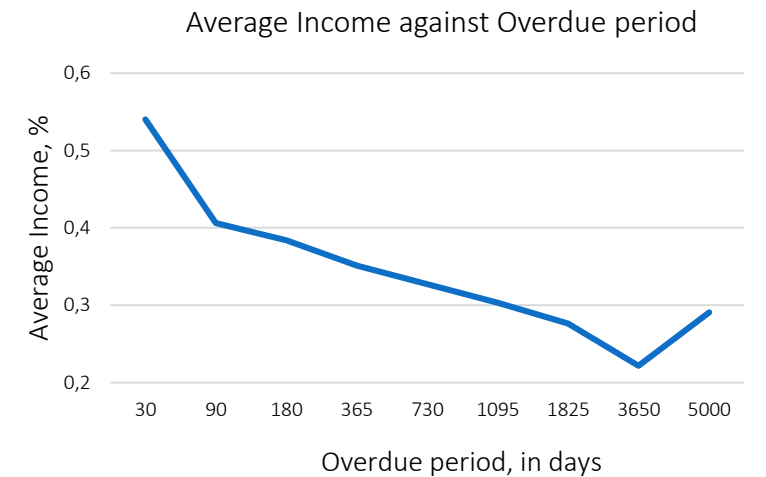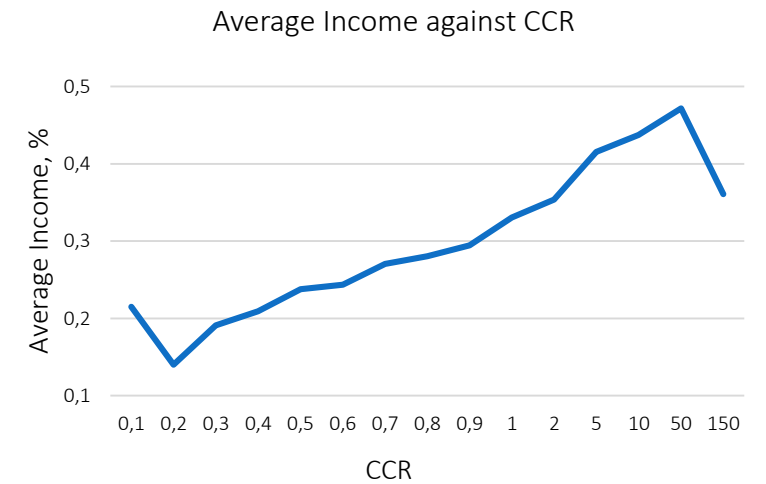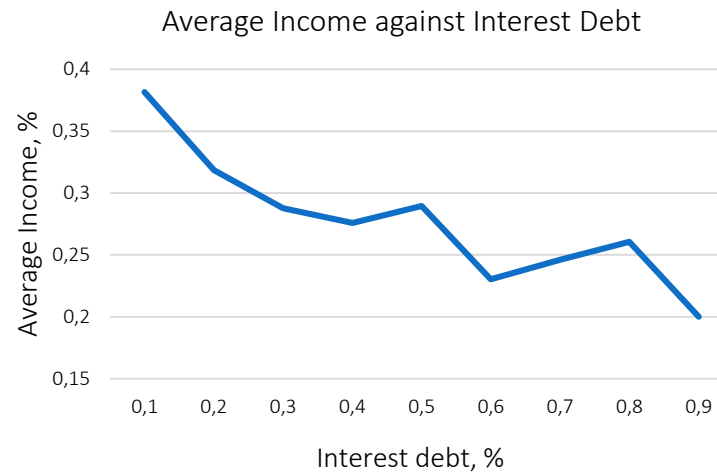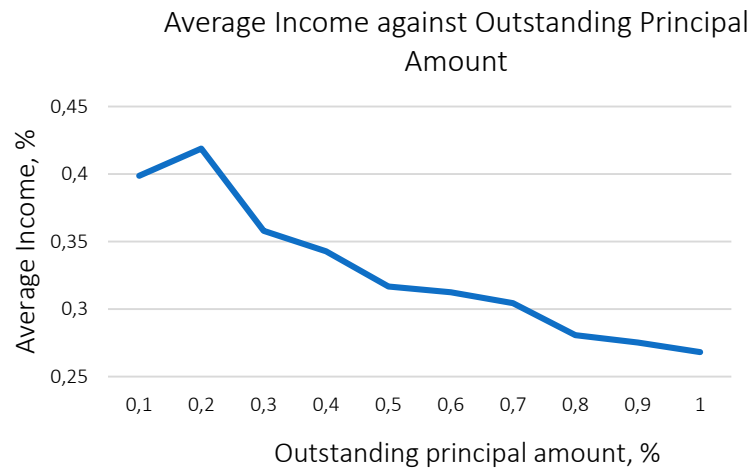
Price Prediction Model[1] – Loans to Individuals (Secured[2]) – Factors

The study identified four factors that significantly impact the loan sale price:

- Interest debt in total debt;

- Outstanding principal amount to the loan amount;

- Collateral coverage ratio (CCR)[3];

- Loan overdue in days.

The graphs demonstrate how Income, % depends on the identified factors.

Average Income against CCR

Average Income against Outstanding Principal Amount

Average Income against Interest Debt

Average Income against Overdue period

16

Price Prediction Model – Loans to Individuals (Secured[1]) – Model

$$P = Total\ debt\ balance \cdot Overdue\ adjustment\ multiplier \cdot Prognosed\ share\ of\ income$$

For prediction of a sale price for standalone loans to individuals, a model combining statistical and econometric features has been developed to estimate the cost of loans based on the four factors

Interest debt

Outstanding principal amount to the loan amount

Collateral Coverage Ratio (CCR)
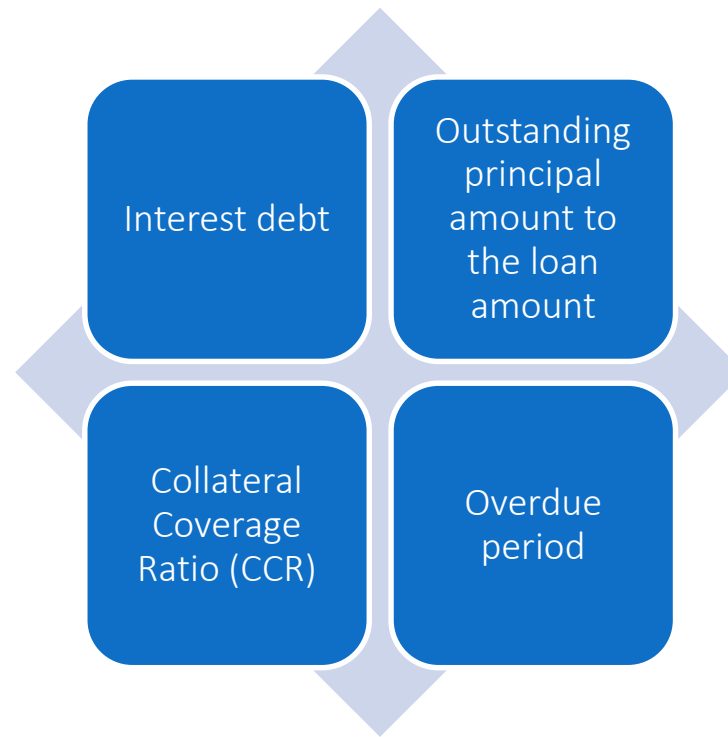
Overdue period

P – prognosed loan income,

Total debt balance – the amount of total loan debt,

Overdue adjustment multiplier – an adjustment multiplier, which is calculated statistically based on the loan overdue period[2],

Prognosed share of income – prognosed percentage of income determined statistically based on Interest debt share, outstanding principal amount, and CCR[2].

[1] As part of this study, a Loans-to-Individuals model was built for secured loans only since the dataset contained no data on unsecured loans to individuals.
[2] Tables with values are provided in the Study document.

Price Prediction Model[1] – Loans to Legal Entities

$$P = Total\ debt\ balance \cdot Portfolio\ adjustment\ multiplier \cdot Forecasted\ share\ of\ income$$
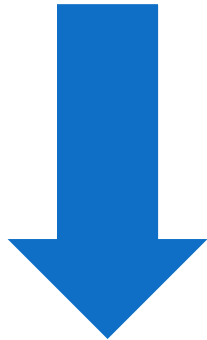
Due to the sample size of loans to legal entities, the use of statistical methods for the analysis is limited as only one main factor that impacts the cost—the size of the loan—may be identified.

Other characteristics of the loan, such as loan currency, overdue period, information about the legal entity, legal factors, potentially have a significant impact on the final sale price. All of them can be taken into account using the *Portfolio adjustment multiplier*.
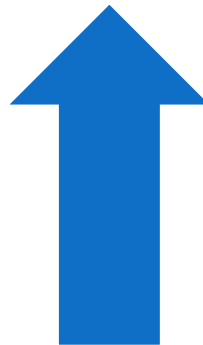
P – prognosed loan income,

Total debt balance – the amount of total loan debt,

Portfolio adjustment multiplier – a loan adjustment coefficient determined by the expert method based on the loan currency, overdue period, existing legal factors[2],

Forecasted share of income – the predicted share of income determined statistically (separately for secured and unsecured loans) based on the size of the loan[2].

[2] Tables with values are provided in the Study document.

[1] Number of observations for the model: 329 (247 secured and 82 unsecured) loans

Price Prediction Model – Pools of Loans to Individuals (Secured)

The model for estimating the price of pools consisting of secured loans to individuals
is based on the model for calculating standalone loans to individuals.

Price of Loan 1 **+** Price of Loan 2 **+** Price of Loan 3 **+** ... **+** Price of Loan N **=** Pool Price
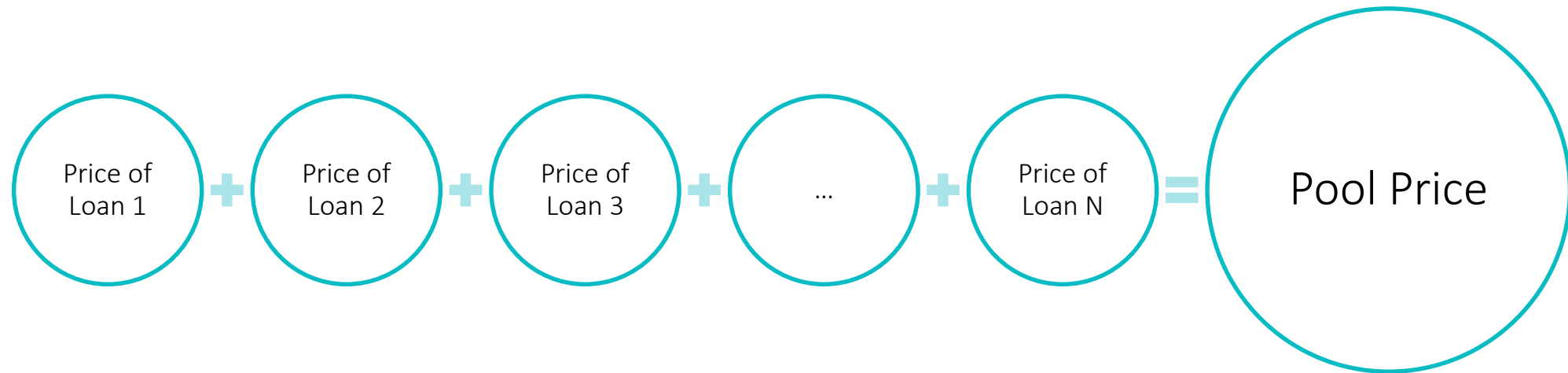
To determine the sale price of such pool, we must calculate the cost
of each loan in the pool using the appropriate model and sum them up.

Price Prediction Model – Pools of Loans to Individuals (Unsecured)

$$P = Total\ debt\ balance \cdot Portfolio\ adjustment\ multiplier \cdot$$
$$Forecasted\ share\ of\ income$$

P – prognosed income of the loan pool,

Total debt balance – the amount of total debt on the pool of loans,

Portfolio adjustment multiplier – an adjustment coefficient for the pool, statistically determined based on the size of the pool and the average overdue period,

Forecasted share of income – the predicted share of income from the pool, determined statistically based on the average loan debts and the overdue period.

Model Factors

Pool size, in UAH

Average loan debt

Average overdue period

Portfolio adjustment multiplier		Average overdue, in days	
		≤ 1095	> 1095
Pool size, in UAH m	≤ 200	1	1
	> 200	1	0.69

Forecasted share of income		Average overdue period, in days	
		≤ 1095	> 1095
Average debt, in UAH	≤ 25 000	1.98%	1.33%
	> 25 000	1.04%	0.76%

Price Prediction Model – Modeling Challenges

The study has identified issues that made it impossible to build models for the following types of assets.

Real estate

The study did not reveal the influence of the appraised value determined by the external appraiser on the sale price of the property at the auction

Insufficient data on certain subtypes to build a high-quality model

Pools of Loans to Legal Entities

Insufficient observations for the analysis and building models

Other
(coins, art, securities)

Atypical bank assets

The share of sales of assets of this type is minor

Due to the variety of assets included in these types, it is impossible to create of homogeneous groups for the analysis and building relevant models

January 2024

Described models and analysis are build on historical data over the period from 2017 to February 2022 placed on the Prozorro.sales platform. All developed methodologies are based on the data on asset sales before the start of the war. As the economic and political situation of the market changes in time, the predictive ability of models may be impaired when applied to data during and after the war. The proposed models are methodological in nature. Models must be periodically reviewed (validated).