



Descriptive Analytics
Framework

Areteans Descriptive Analytics Framework in Pega(aDAFP)

An intelligent and powerful AI framework to translate insights into growth.



Global Elite
Partner



Areteans Descriptive Analytics Framework in Pega(aDAFP)

An intelligent and powerful AI framework to translate insights into growth.

What is aDAFP?

Areteans Descriptive Analytics Framework in Pega(aDAFP) in the space Artificial Intelligence, powered by Inferential Statistics, is a prototype solution, refers to analyze, summarize, visualize business Data and represent the historical discoveries and outcomes of Pega customer dataset, derived from a sample or population, in Feature Engineering and Preprocessing phase of Data Science project Life cycle. This framework will also facilitates to explore the statistical information and pattern recognition from customer behavioral, demographic or transactional past data and to built the state-of-the-art Statistical AI Model.

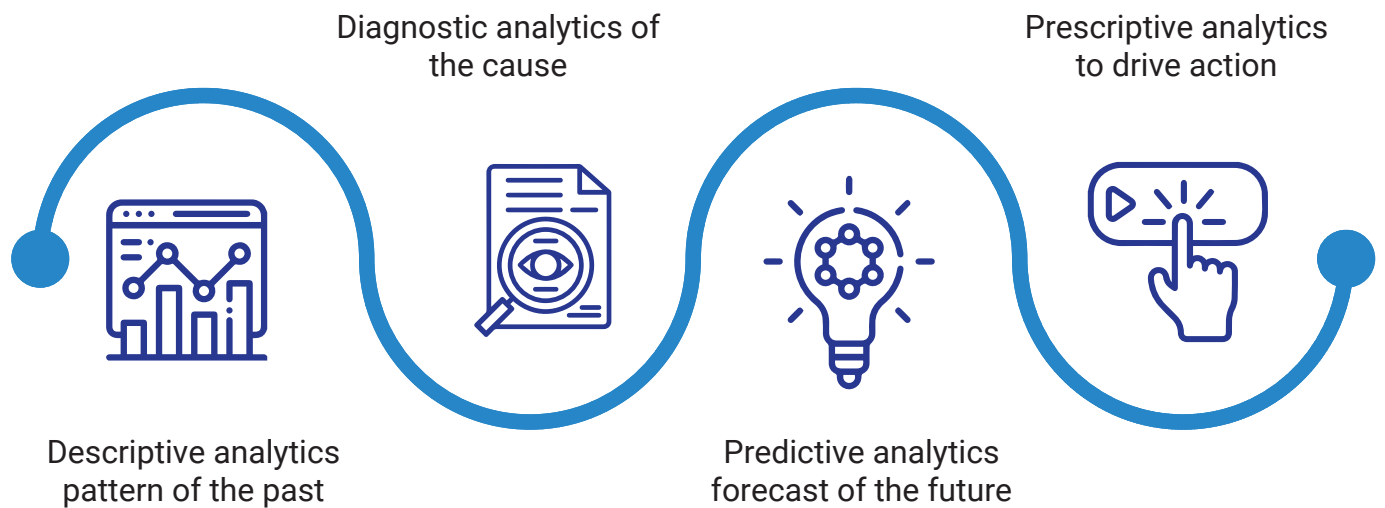


Why is aDAFP?

- Helps customers to enable business data visualization to implement business decisions – complementing Pega's OOTB AI capability which most of the CRM and Data Science platforms provide.
- Powerful aid provides statistical information about variables and highlights potential relationships between explanatory variables and response variables. It also condenses the characteristics of a sample from a population such as predictors mean, median, standard deviation, or dispersion – and helps to develop high-performance AI Algorithms with suitable Predictor and other configurations.
- Determine the pattern of the data and helps eliminate the presence of noise before AI Model develops.



Need for Descriptive Analytics in Pega Decisioning Engine



Key differentiators

- In Pega's Prediction studio, suitable data preprocessing such as constant data, blank-NA-missing data, outlier data, imbalance data, redundant data, dispersion analysis, etc. aren't being fully explored in the Predictive Model creation. Sometimes a data scientist utilizes external AI tools/platforms to preprocess the data/predictor and import it into Pega's Prediction studio.
- Pega Model Development wizard requires a developer to select some specifications like Analyzing dummy variables, Type of sampling (Uniform, Stratified, etc.), Creation of a New Derived Predictor, and setting a threshold of performance which are not immediately available and hence needs external analysis or domain expertise of data to build it correctly.
- Pega's Adaptive Model is driven by Naïve Bayes Algorithm(before 2022) only and this Algorithm works perfectly only when there is no/min intercorrelation among predictors. There is hardly any straightaway wizard to analyze this before building Pega's Adaptive Model.'





Unique features

Pre-process the data and predictors on a unified platform of Pega

Recognize the patterns of trends and anomalies which are hidden in raw data

Prepare, extract and generate relevant data for modeling in Pega

Visualize important statistical KPIs and metrics of historical performances

Leverage plugged-in Pega customer decision hub portal

Support future enhancements with customer requirements/data with robust and scalable components

Embed in any Pega 8x engine with simple JavaScript/CSS or open-sourced JS APIs

Solution capabilities

Heatmap Analysis

Provides AI Heatmap Plot of data to check the statistical interrelation among explanatory variables with response variables where color intensity represents the depth of the relationship

Box and Whisker Analysis

- **Minimum Score** - The lowest score, excluding outliers (shown at the end of the left whisker)
- **Lower Quartile** – 25% of scores fall below the lower quartile value (also known as the first quartile)
- **Median** - The median marks the mid-point of the data and is shown by the line that divides the box into two parts
- **Upper Quartile** – 75% percent of the scores fall below the upper quartile value (also known as the third quartile). Thus, 25% of the data are above this value
- **Maximum Score** - The highest score, excluding outliers (shown at the end of the right whisker)
- **Whiskers** - The upper and lower whiskers represent scores outside the middle 50% (i.e. the lower 25% of scores and the upper 25% of scores)
- **The Interquartile Range (or IQR)** - This is the box plot showing the middle 50% of scores (i.e., the range between the 25th and 75th percentile)
- **Outlier detection** of the dataset for each predictor
- **Visualize the dispersion** and distribution of data

Histogram Analysis

A histogram plot of predictors is to check the spread, dispersion, imbalance, and skewness of data and hence can easily observe the pattern of the distribution i.e. Gaussian-Normal Distribution, or exponential Distribution or Bi-modal Distribution, etc.

Basic Feature Engineering

Statistical Data Preprocessing Analytics for feature engineering with detailed analysis, for Numerical Predictors, of

- Mean
- Count/Frequency of Data
- Standard Deviation
- Minimum and Maximum value of Data
- 25-50-75 Percentile

Imbalance and Bias effect of Target Variable

Detect Class Imbalance of the target Class(for Classification Type of Problem) and hence need to take appropriate transformations or measures to remove Biasing of target





Global Elite
Partner

