# H2O-3 Experiment

Generated by: H2O\_from\_python\_mathanraj\_sp3xcl

Generated on: 2021-10-01 14:08

[H2O-3 Experiment 1](#_Toc38398026)

[Experiment Overview 1](#_Toc38398027)

[Data Overview 1](#_Toc32075692)

[Validation Strategy 1](#_Toc32075693)

[Feature Importance 1](#_Toc38398030)

[Final Model 1](#_Toc38398031)

[Alternative Models 1](#_Toc38398032)

[Partial Dependence Plots 2](#_Toc38398033)

[Model Reproducibility 2](#_Toc38398034)

[Appendix 2](#_Toc38398035)

## Experiment Overview

H2O-3 built a Gradient Boosting Machine to predict DEFAULT\_PAYMENT\_NEXT\_MONTH given 23 original features from the input dataset. This classification experiment completed in 4 seconds (0:00:04).

### Performance

|  |  |
| --- | --- |
| **Dataset** | **auc** |
| Validation Data | 0.801 |
| Test Data | NA |

### System Specifications

|  |  |
| --- | --- |
| **Attribute** | **Value** |
| H2O cluster uptime: | 59 secs |
| H2O cluster timezone: | Asia/Colombo |
| H2O data parsing timezone: | UTC |
| H2O cluster version: | 3.32.1.7 |
| H2O cluster version age: | 28 days, 19 hours and 12 minutes |
| H2O cluster name: | H2O\_from\_python\_mathanraj\_sp3xcl |
| H2O cluster total nodes: | 1 |
| H2O cluster free memory: | 3.229 Gb |
| H2O cluster total cores: | 8 |
| H2O cluster allowed cores: | 8 |
| H2O cluster status: | locked, healthy |
| H2O connection url: | http://127.0.0.1:54321 |
| H2O connection proxy: | {'http': None, 'https': None} |
| H2O internal security: | False |
| H2O API Extensions: | Amazon S3, XGBoost, Algos, AutoML, Core V3, TargetEncoder, Core V4 |
| Python version: | 3.7.10 final |

### Versions

|  |  |
| --- | --- |
| **Package** | **Version Number** |
| H2O-3 | 3.32.1.7 |

## Data Overview

This section provides information on the datasets used for the experiment.

|  |  |  |
| --- | --- | --- |
| **data** | **rows** | **cols** |
| train | 23,999 | 25 |
| validation | 6,000 | 25 |

### Training Data

The training data consists of both numeric and categorical columns.

The summary of the columns is shown below:

#### Numeric Columns

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **name** | **min** | **mean** | **max** | **std** |
| ID | 1 | 1.2e+04 | 2.4e+04 | 6928 |
| LIMIT\_BAL | 1e+04 | 1.655e+05 | 1e+06 | 1.291e+05 |
| AGE | 21 | 35.38 | 79 | 9.271 |
| PAY\_0 | -2 | -0.001667 | 8 | 1.127 |
| PAY\_2 | -2 | -0.1235 | 8 | 1.201 |
| PAY\_3 | -2 | -0.1548 | 8 | 1.204 |
| PAY\_4 | -2 | -0.2117 | 8 | 1.167 |
| PAY\_5 | -2 | -0.2529 | 8 | 1.137 |
| PAY\_6 | -2 | -0.278 | 8 | 1.158 |
| BILL\_AMT1 | -1.656e+05 | 5.06e+04 | 9.645e+05 | 7.265e+04 |
| BILL\_AMT2 | -6.978e+04 | 4.865e+04 | 9.839e+05 | 7.037e+04 |
| BILL\_AMT3 | -1.573e+05 | 4.637e+04 | 1.664e+06 | 6.819e+04 |
| BILL\_AMT4 | -1.7e+05 | 4.237e+04 | 8.916e+05 | 6.307e+04 |
| BILL\_AMT5 | -8.133e+04 | 4e+04 | 9.272e+05 | 6.035e+04 |
| BILL\_AMT6 | -3.396e+05 | 3.857e+04 | 9.617e+05 | 5.916e+04 |
| PAY\_AMT1 | 0 | 5543 | 5.05e+05 | 1.507e+04 |
| PAY\_AMT2 | 0 | 5816 | 1.684e+06 | 2.08e+04 |
| PAY\_AMT3 | 0 | 4969 | 8.96e+05 | 1.61e+04 |
| PAY\_AMT4 | 0 | 4744 | 4.97e+05 | 1.488e+04 |
| PAY\_AMT5 | 0 | 4784 | 4.18e+05 | 1.527e+04 |
| PAY\_AMT6 | 0 | 5190 | 5.287e+05 | 1.763e+04 |

#### Categorical Columns

|  |  |  |  |
| --- | --- | --- | --- |
| **name** | **unique** | **top** | **freq of top value** |
| SEX | 2 | female | 15078 |
| EDUCATION | 4 | university | 11360 |
| MARRIAGE | 4 | single | 12876 |
| DEFAULT\_PAYMENT\_NEXT\_MONTH | 2 | 0 | 18630 |

#### Shifts Detected

H2O-3 can perform shift detection between the training, validation, and testing datasets. It does this by training a binomial model to predict which dataset a record belongs to. For example, it may find that it is able to separate the training and testing data with an AUC of 0.8 using only the column: C1 as the predictor. This indicates that there is some sort of drift in the distribution of C1 between the training and testing data.

For this experiment, H2O-3 checked the train and validation data for any shift in distributions but found none. This indicates that all the predictors/columns in the train and validation data are from the same distribution.

## Validation Strategy

The model’s performance is evaluated using a validation dataset with shape (6000, 25).

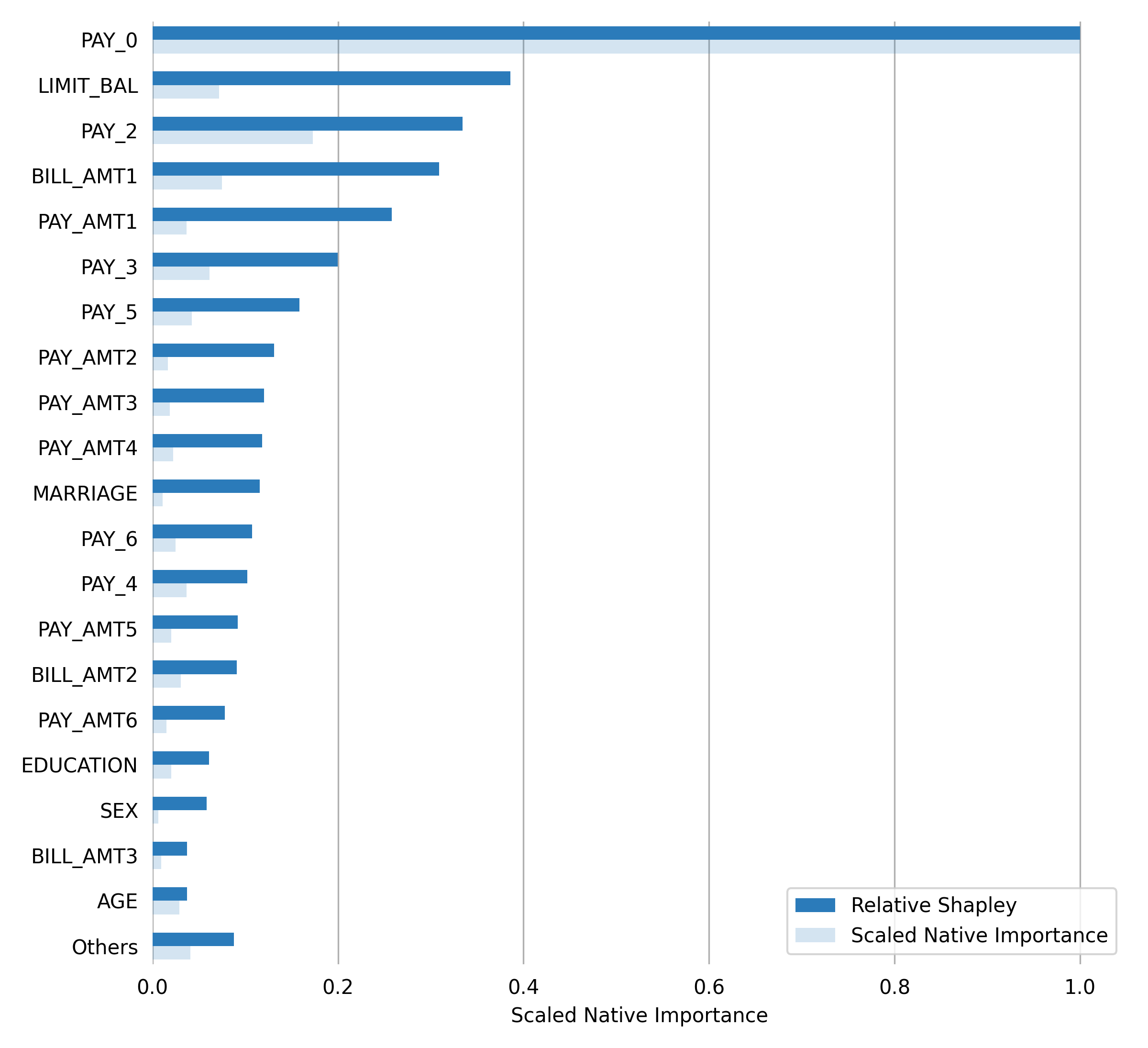
Early stopping ends model training when the selected “stopping metric” does not improve for a specified number of training rounds, based on a simple moving average. For this model, early stopping is disabled during training.

## Feature Importance

H2O-3 models provide built-in variable importance (Native Importance) and can provide Shapley Importance for supported algorithms.

* **Native Importance:** Model-specific variable importance calculated with H2O-3’s varimp() function (H2O-3 documentation details [here](https://h2o-release.s3.amazonaws.com/h2o/rel-zipf/7/docs-website/h2o-docs/variable-importance.html)).
* **Scaled Native Importance:** Native Importance scaled between 0 and 1.
* **Shapley:** The mean absolute Shapley value of a feature, using TreeSHAP (SHAP documentation details [here](https://shap.readthedocs.io/en/latest/)).
* **Relative Shapley:** The feature’s mean absolute Shapley value divided by the largest Shapley value.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Feature** | **Native Importance** | **Scaled Native Importance** | **Shapley** | **Relative Shapley** |
| 0 | PAY\_0 | 3278.1111 | 1.0 | 0.4045 | 1.0 |
| 1 | LIMIT\_BAL | 235.791 | 0.0719 | 0.1561 | 0.3859 |
| 2 | PAY\_2 | 567.0049 | 0.173 | 0.1352 | 0.3343 |
| 3 | BILL\_AMT1 | 245.1739 | 0.0748 | 0.125 | 0.3089 |
| 4 | PAY\_AMT1 | 120.3364 | 0.0367 | 0.1043 | 0.2579 |
| 5 | PAY\_3 | 201.1064 | 0.0613 | 0.0807 | 0.1996 |
| 6 | PAY\_5 | 139.495 | 0.0426 | 0.0642 | 0.1587 |
| 7 | PAY\_AMT2 | 53.8554 | 0.0164 | 0.053 | 0.1311 |
| 8 | PAY\_AMT3 | 60.9946 | 0.0186 | 0.0486 | 0.1201 |
| 9 | PAY\_AMT4 | 73.9498 | 0.0226 | 0.0479 | 0.1185 |
| 10 | MARRIAGE | 35.4167 | 0.0108 | 0.0468 | 0.1157 |
| 11 | PAY\_6 | 81.0877 | 0.0247 | 0.0434 | 0.1073 |
| 12 | PAY\_4 | 120.4323 | 0.0367 | 0.0414 | 0.1024 |
| 13 | PAY\_AMT5 | 66.5133 | 0.0203 | 0.0372 | 0.092 |
| 14 | BILL\_AMT2 | 99.895 | 0.0305 | 0.0367 | 0.0908 |
| 15 | PAY\_AMT6 | 49.6303 | 0.0151 | 0.0315 | 0.0779 |
| 16 | EDUCATION | 67.2093 | 0.0205 | 0.0246 | 0.0609 |
| 17 | SEX | 20.7221 | 0.0063 | 0.0237 | 0.0587 |
| 18 | BILL\_AMT3 | 30.2772 | 0.0092 | 0.0151 | 0.0372 |
| 19 | AGE | 95.7084 | 0.0292 | 0.015 | 0.037 |
| 20 | Others | 134.0352 | 0.0409 | 0.0355 | 0.0878 |



## Final Model

|  |  |
| --- | --- |
| **Parameters** | **Values** |
| model\_id | gbm\_model |
| auc\_type | AUTO |
| balance\_classes | False |
| categorical\_encoding | Enum |
| class\_sampling\_factors | None |
| col\_sample\_rate | 1.0 |
| col\_sample\_rate\_change\_per\_level | 1.0 |
| col\_sample\_rate\_per\_tree | 1.0 |
| distribution | bernoulli |
| fold\_assignment | None |
| fold\_column | None |
| histogram\_type | UniformAdaptive |
| learn\_rate | 0.1 |
| learn\_rate\_annealing | 1.0 |
| max\_abs\_leafnode\_pred | 1.7976931348623157e+308 |
| max\_after\_balance\_size | 5.0 |
| max\_depth | 5 |
| max\_runtime\_secs | 0.0 |
| min\_rows | 10.0 |
| min\_split\_improvement | 1e-05 |
| nbins | 20 |
| nbins\_cats | 1024 |
| nbins\_top\_level | 1024 |
| ntrees | 50 |
| offset\_column | None |
| pred\_noise\_bandwidth | 0.0 |
| response\_column | DEFAULT\_PAYMENT\_NEXT\_MONTH |
| sample\_rate | 1.0 |
| sample\_rate\_per\_class | None |
| seed | 1234 |
| stopping\_metric | None |
| stopping\_rounds | 0 |
| stopping\_tolerance | 0.001 |
| weights\_column | None |

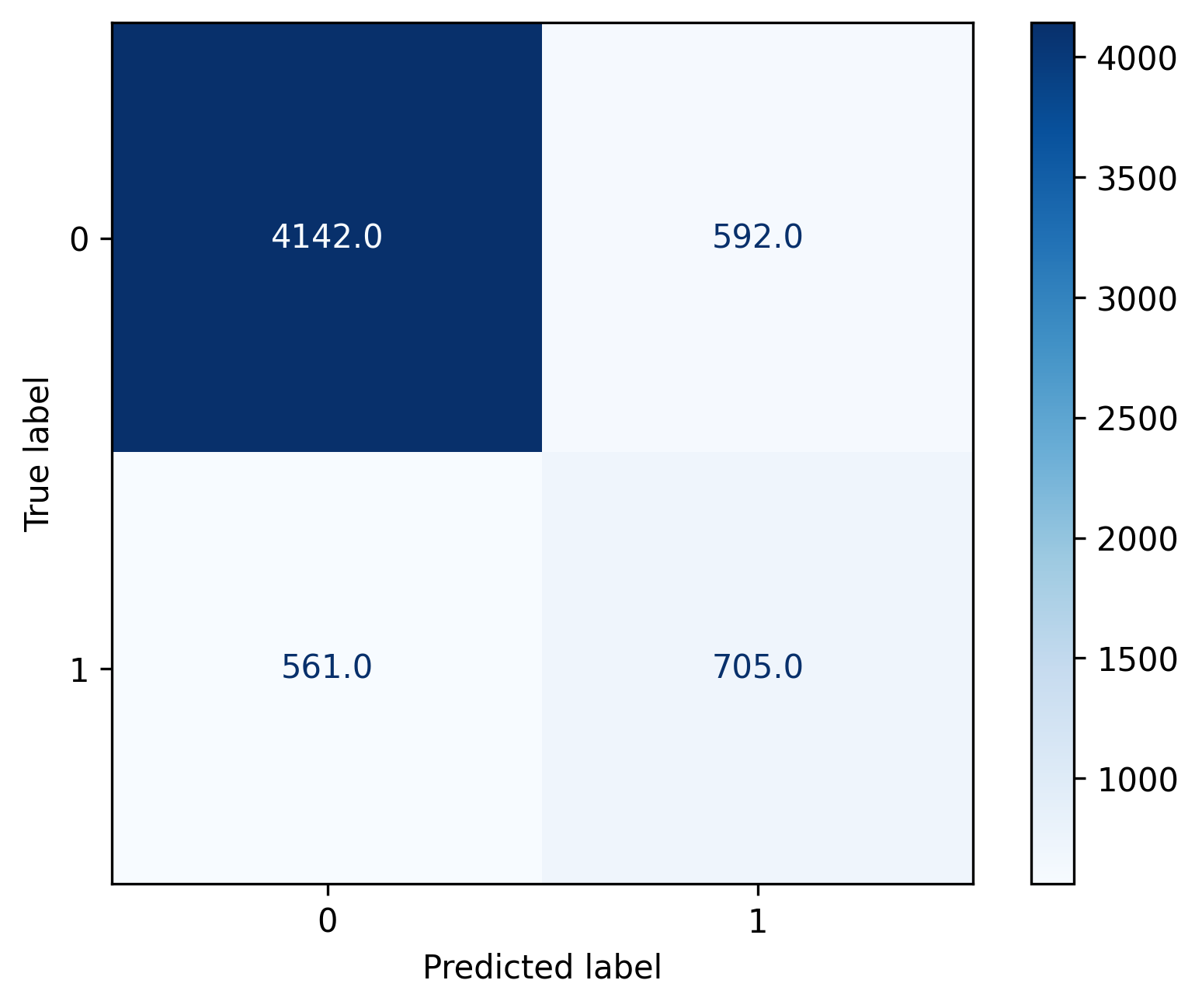
**Performance of Final Model**

|  |  |  |
| --- | --- | --- |
| **scorer** | **training** | **validation** |
| AUC | 0.8116 | 0.8007 |
| ACCURACY | 0.831 | 0.833 |
| F1 | 0.5727 | 0.5501 |
| MCC | 0.4599 | 0.442 |
| LOGLOSS | 0.4087 | 0.4065 |

**Validation Confusion Matrix**

The confusion matrix shows how many observations the model correctly classified and misclassified. The first column contains the actual class labels; the first row contains the predicted class labels.

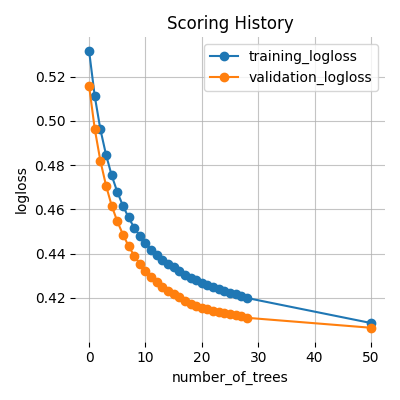
A positive prediction label (e.g., 1, True, or the second label in lexicographical order), is assigned to all observations where the predicted probability is greater than or equal to 0.2688 (the threshold for the highest F1 score on the validation dataset).



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Labels** | **0** | **1** | **Error** | **Rate** |
| 0 | 4142.0 | 592.0 | 0.1251 | (592.0/4734.0) |
| 1 | 561.0 | 705.0 | 0.4431 | (561.0/1266.0) |
| Total | 4703.0 | 1297.0 | 0.1922 | (1153.0/6000.0) |

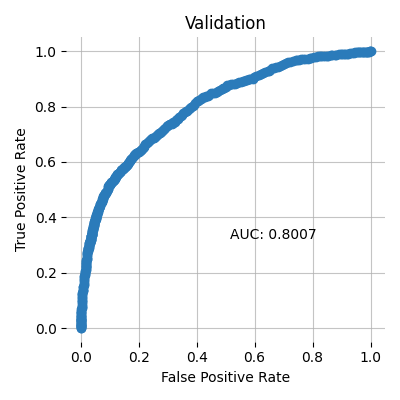
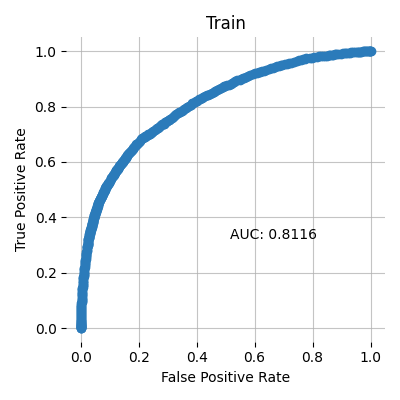
**Scoring History**

The scoring history plot shows a model's performance at each iteration point. Typically, the performance will be worse at the beginning (the left side of the graph) and then improve as the model training completes and accuracy improves.



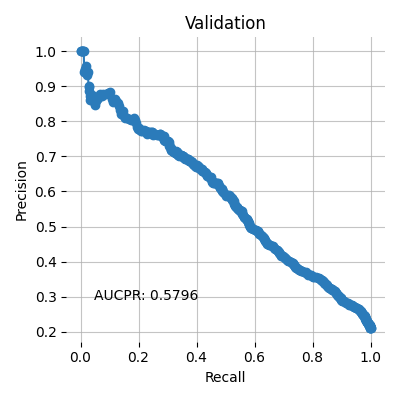
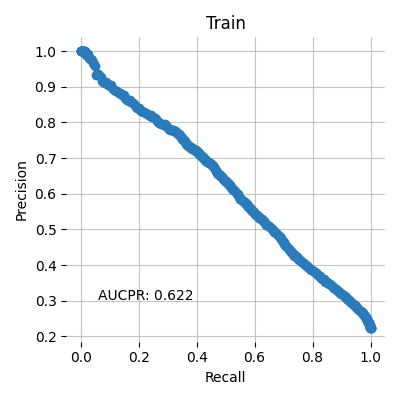
**Receiver Operating Characteristic Curve**

This plot shows the Receiver Operating Characteristic Curve. The area under this curve is called the AUC. The True Positive Rate (TPR) is the relative fraction of correct positive predictions, and the False Positive Rate (FPR) is the relative fraction of incorrect positive corrections. Each point corresponds to a classification threshold (e.g., YES if probability >= 0.3 else NO). For each threshold, there is a unique confusion matrix that represents the balance between TPR and FPR. In general, the most useful operating points are in the top left corner.



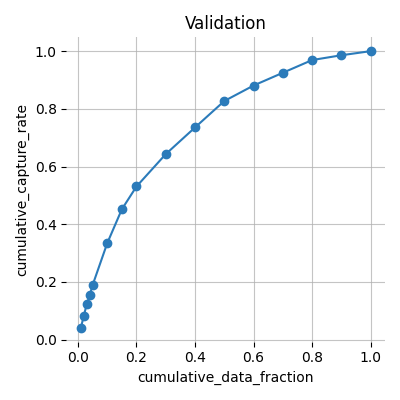
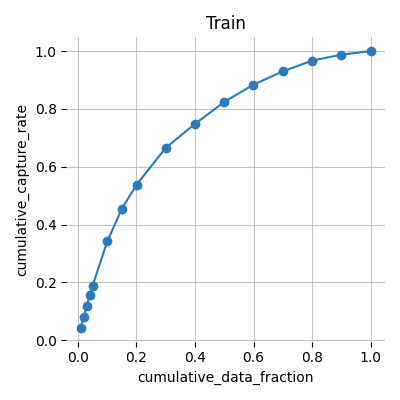
**Precision-Recall Curve**

This model metric is used to evaluate how well a binary classification model is able to distinguish between precision recall pairs or points. These values are obtained using different thresholds on a probabilistic or other continuous-output classifier.



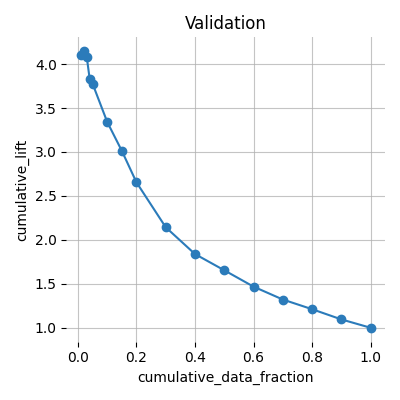
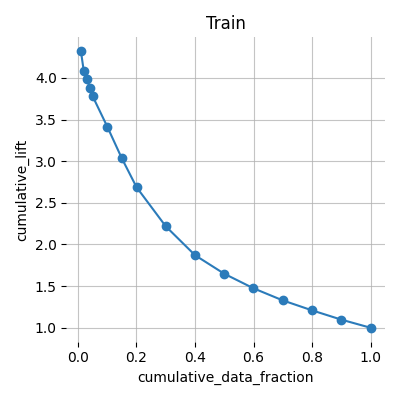
**Cumulative Gain**

This plot shows the cumulative gains. For example, "What fraction of all observations of the positive target class are in the top predicted 1%, 2%, 10%, etc. (cumulative)?" The gains at 100% are 1.0 by definition.



**Cumulative Lift**

This chart shows the cumulative lift. For example, "How many times more observations of the positive target class are in the top predicted 1%, 2%, 10%, etc. (cumulative) compared to selecting observations randomly?" By definition, the Lift at 100% is 1.0.



**Population Stability Index (PSI)**

Population Stability Index is a statistic used to describe a variable’s distribution shift. It can measure the shift between the training dataset’s model score distribution and any other given dataset (i.e. validation or test dataset).

A PSI value lower than 0.10 indicates a small shift in the model predictions, a value between 0.10 and 0.25 indicates a moderate shift, and a value greater than 0.25 indicates a strong shift. Strong shift values can indicate that the model trained on the training dataset might not be suitable for the provided validation or test datasets.

**Summary PSI table**

|  |  |
| --- | --- |
| **Dataset** | **PSI** |
| Validation | 0.0052 |

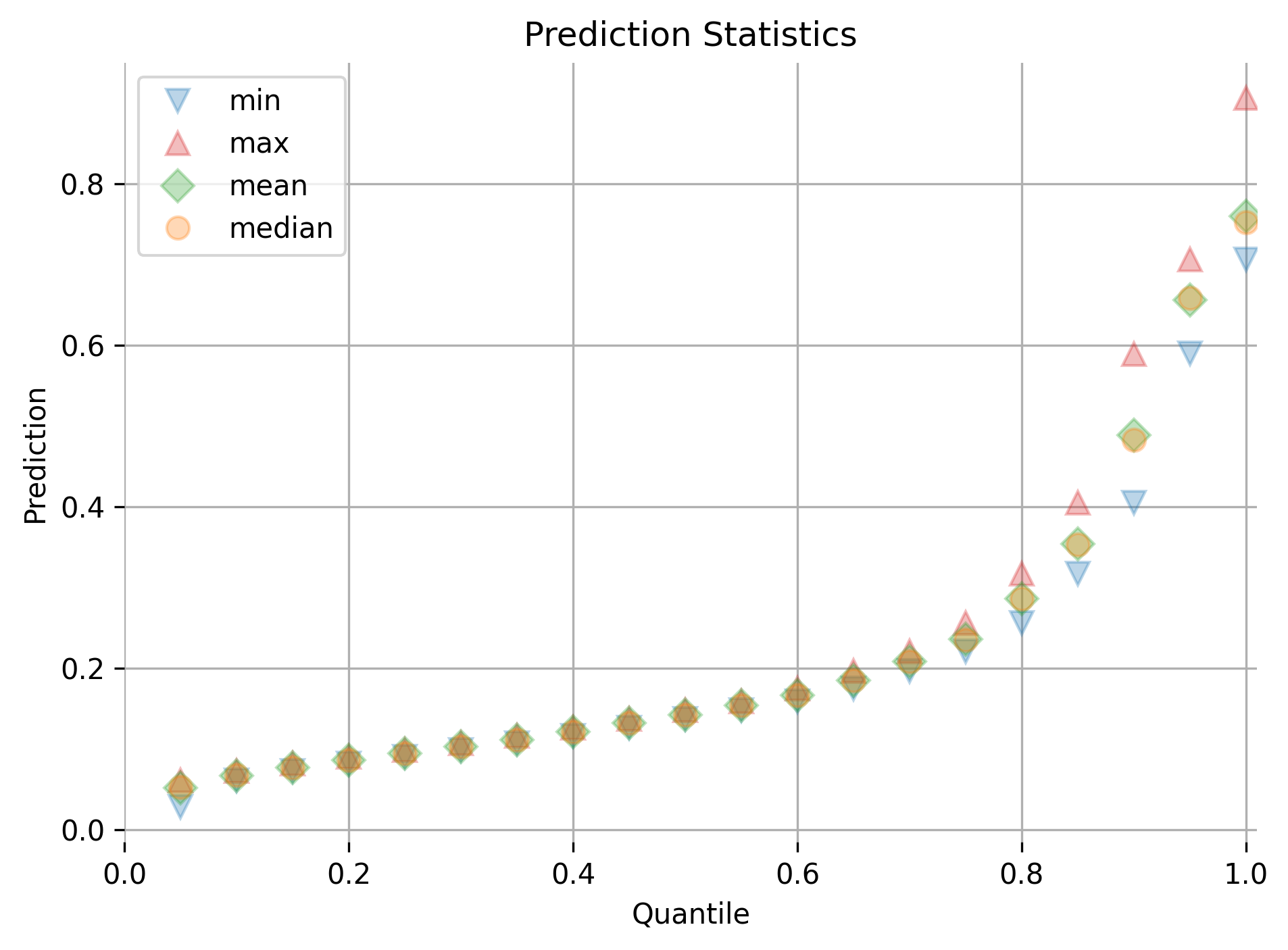
Details on the PSI calculations can be found in the Appendix.

**Prediction statistics**

The following tables and plots show the min, max, mean, and median quantile prediction values for each dataset split. Note: values are rounded to the fourth decimal place. For example, .000025 and .000010 would both appear as 0.0.

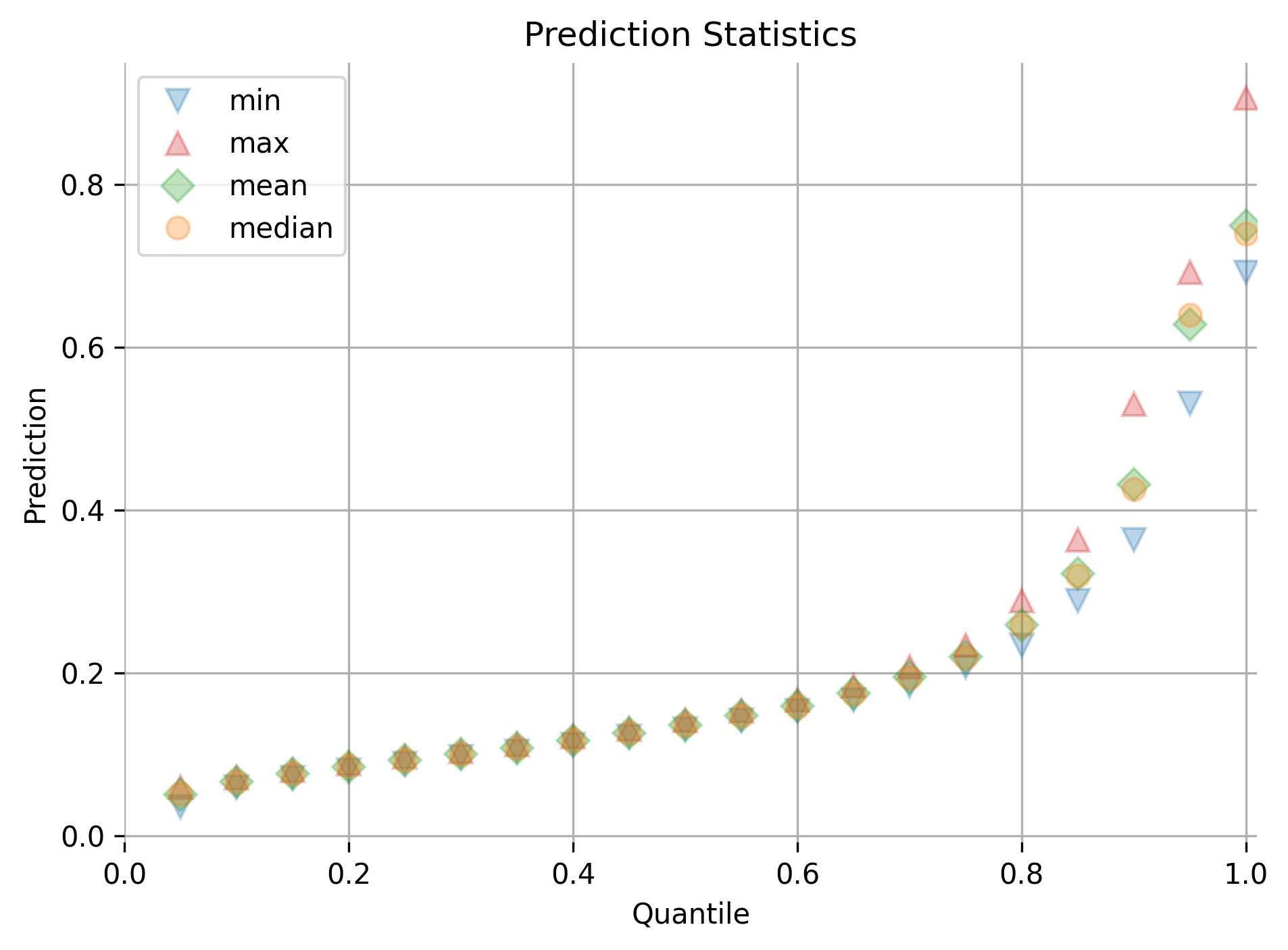
**Train**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **quantile** | **min** | **max** | **mean** | **median** |
| 0.05 | 0.0283 | 0.0607 | 0.0516 | 0.0527 |
| 0.1 | 0.0607 | 0.0721 | 0.0668 | 0.0671 |
| 0.15 | 0.0721 | 0.0815 | 0.0769 | 0.077 |
| 0.2 | 0.0815 | 0.0902 | 0.0858 | 0.0857 |
| 0.25 | 0.0902 | 0.0982 | 0.0942 | 0.0941 |
| 0.3 | 0.0982 | 0.1066 | 0.1024 | 0.1023 |
| 0.35 | 0.1066 | 0.1159 | 0.1112 | 0.1113 |
| 0.4 | 0.1159 | 0.1264 | 0.1211 | 0.121 |
| 0.45 | 0.1264 | 0.1369 | 0.1316 | 0.1315 |
| 0.5 | 0.1369 | 0.1475 | 0.1421 | 0.142 |
| 0.55 | 0.1475 | 0.159 | 0.1532 | 0.1534 |
| 0.6 | 0.159 | 0.1745 | 0.1664 | 0.1662 |
| 0.65 | 0.1745 | 0.196 | 0.1848 | 0.1849 |
| 0.7 | 0.1961 | 0.2204 | 0.2083 | 0.2084 |
| 0.75 | 0.2204 | 0.2559 | 0.2356 | 0.2349 |
| 0.8 | 0.2559 | 0.3169 | 0.2857 | 0.2856 |
| 0.85 | 0.3169 | 0.4048 | 0.354 | 0.3527 |
| 0.9 | 0.4049 | 0.5891 | 0.4888 | 0.4822 |
| 0.95 | 0.5893 | 0.7063 | 0.6561 | 0.6587 |
| 1.0 | 0.7063 | 0.9057 | 0.7596 | 0.7519 |



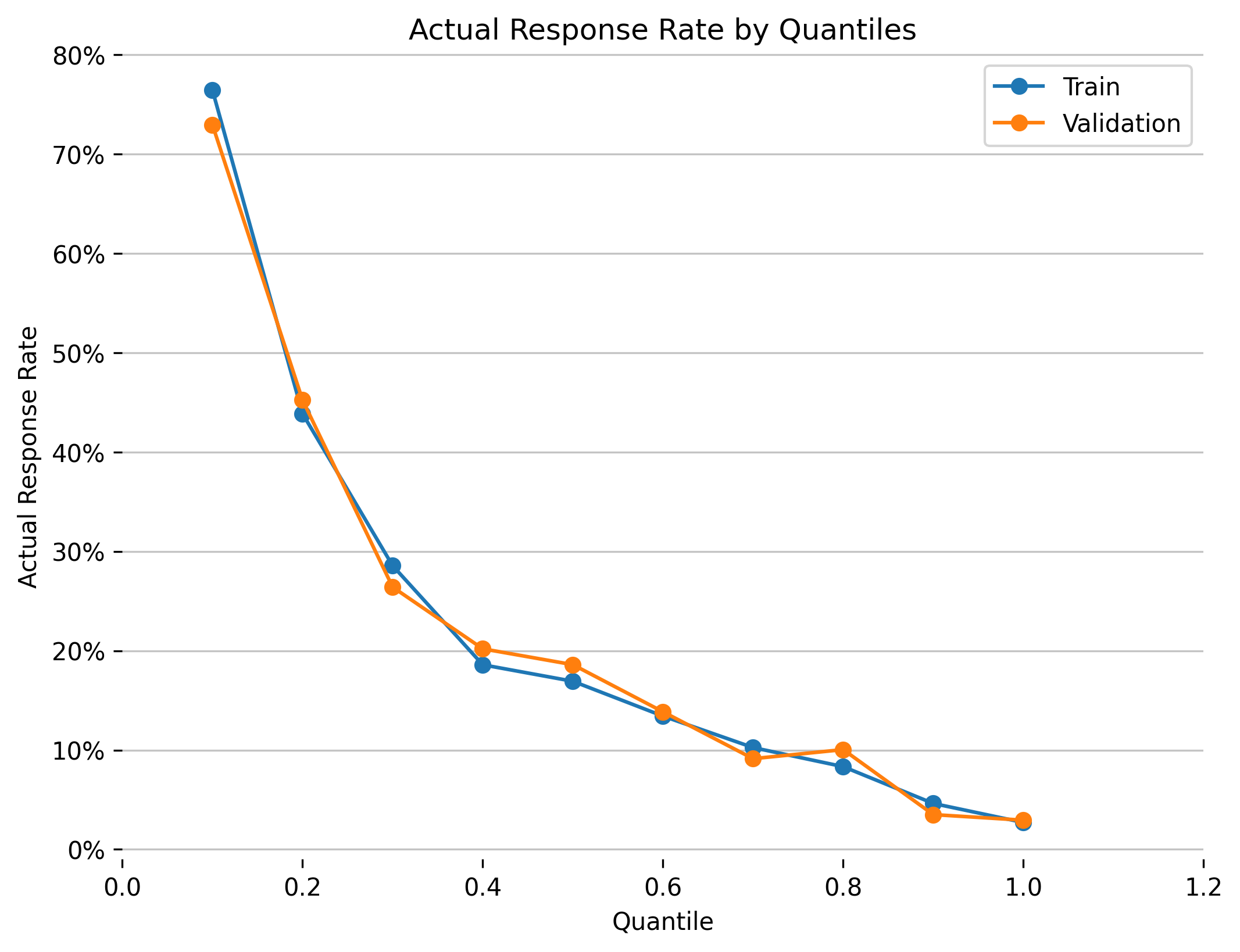
**Validation**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **quantile** | **min** | **max** | **mean** | **median** |
| 0.05 | 0.0361 | 0.0599 | 0.0509 | 0.0518 |
| 0.1 | 0.06 | 0.0715 | 0.0662 | 0.0667 |
| 0.15 | 0.0715 | 0.0806 | 0.0762 | 0.0764 |
| 0.2 | 0.0806 | 0.0891 | 0.085 | 0.0853 |
| 0.25 | 0.0891 | 0.0969 | 0.0931 | 0.0931 |
| 0.3 | 0.0969 | 0.1042 | 0.1004 | 0.1004 |
| 0.35 | 0.1042 | 0.1121 | 0.1079 | 0.1079 |
| 0.4 | 0.1122 | 0.1219 | 0.1171 | 0.1174 |
| 0.45 | 0.1219 | 0.1308 | 0.1264 | 0.1264 |
| 0.5 | 0.1309 | 0.1418 | 0.1361 | 0.1359 |
| 0.55 | 0.1418 | 0.1536 | 0.1474 | 0.1477 |
| 0.6 | 0.1536 | 0.1667 | 0.1597 | 0.159 |
| 0.65 | 0.1667 | 0.1847 | 0.175 | 0.1748 |
| 0.7 | 0.1847 | 0.2072 | 0.1947 | 0.1944 |
| 0.75 | 0.2072 | 0.2339 | 0.2198 | 0.2189 |
| 0.8 | 0.234 | 0.2885 | 0.2592 | 0.2578 |
| 0.85 | 0.2886 | 0.3635 | 0.322 | 0.3195 |
| 0.9 | 0.3637 | 0.5304 | 0.4312 | 0.4256 |
| 0.95 | 0.5308 | 0.6917 | 0.6283 | 0.6397 |
| 1.0 | 0.6918 | 0.9056 | 0.7495 | 0.7396 |



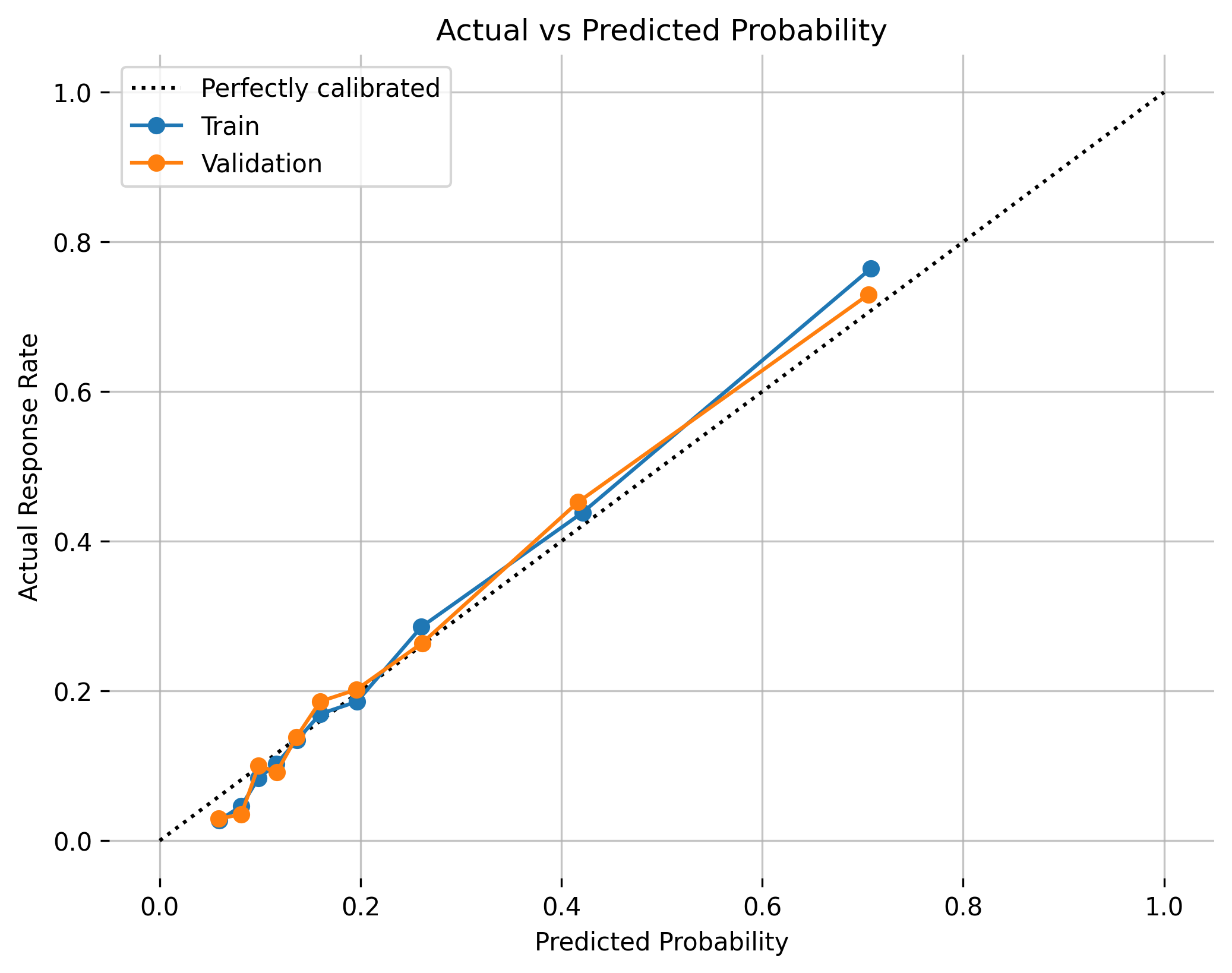
**Quantile Response Rates**

The response rate, for a given quantile, is equal to the number of positive-labeled data points divided by the total number of data points. Quantiles are sorted in decreasing order.



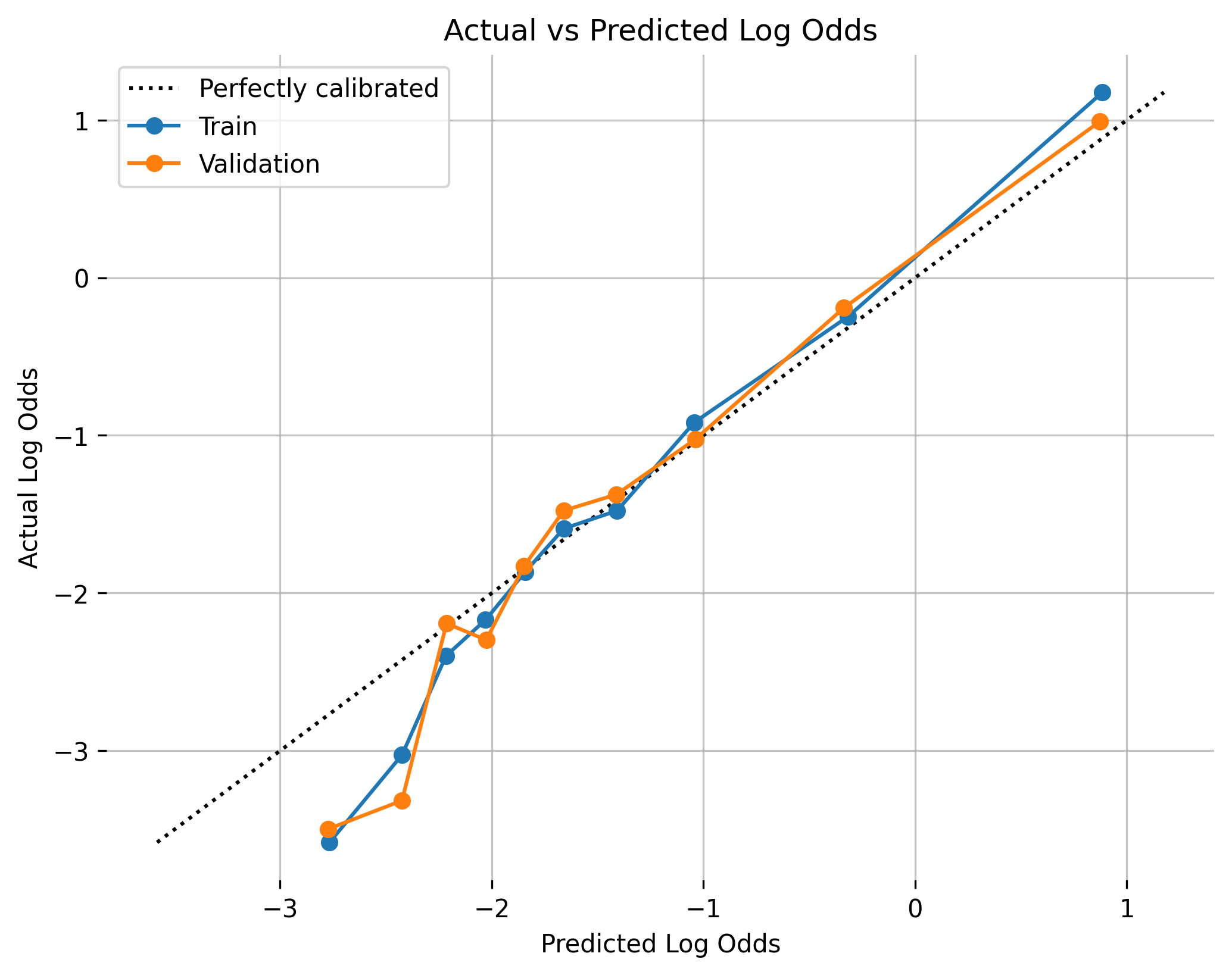
**Actual vs Predicted Probabilities**

This plot shows the alignment between the predicted and the actual probabilities. The predicted probabilities are binned into quantiles. For each, bin the average predicted value and the actual response rate (i.e., the number positive-labeled records divided by the total number of records within each bin) is calculated.



**Actual vs Predicted Log Odds**

This plot shows the alignment between the predicted and the actual probabilities within the log odds space. In this case, the log odds are the log transformation of the probability of a positive record divided by the probability of a negative record.



Details on the quantile-based plots’ calculations can be found in the Appendix.

## Alternative Models

Alternative models were not provided by the user.

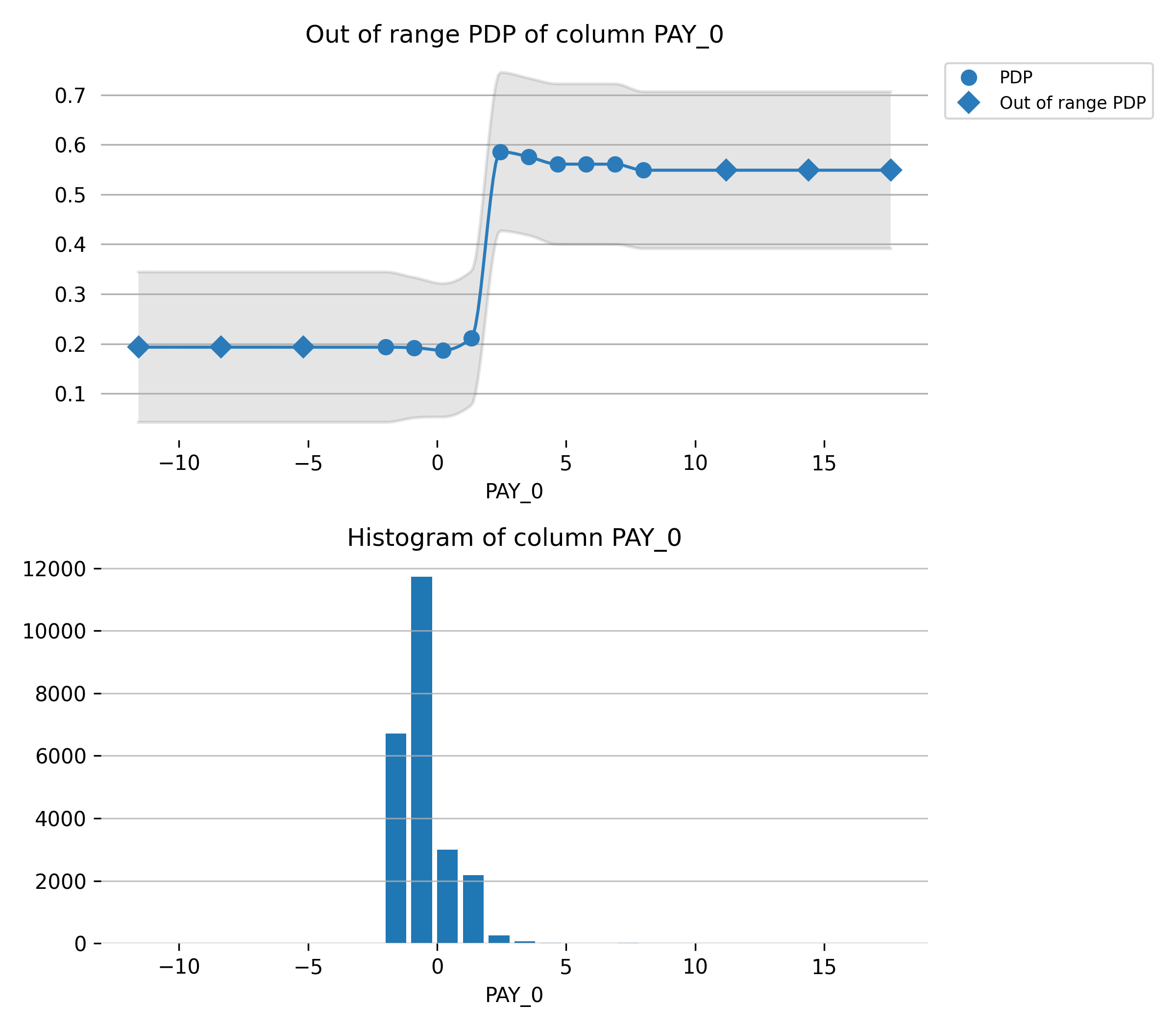
## Partial Dependence Plots

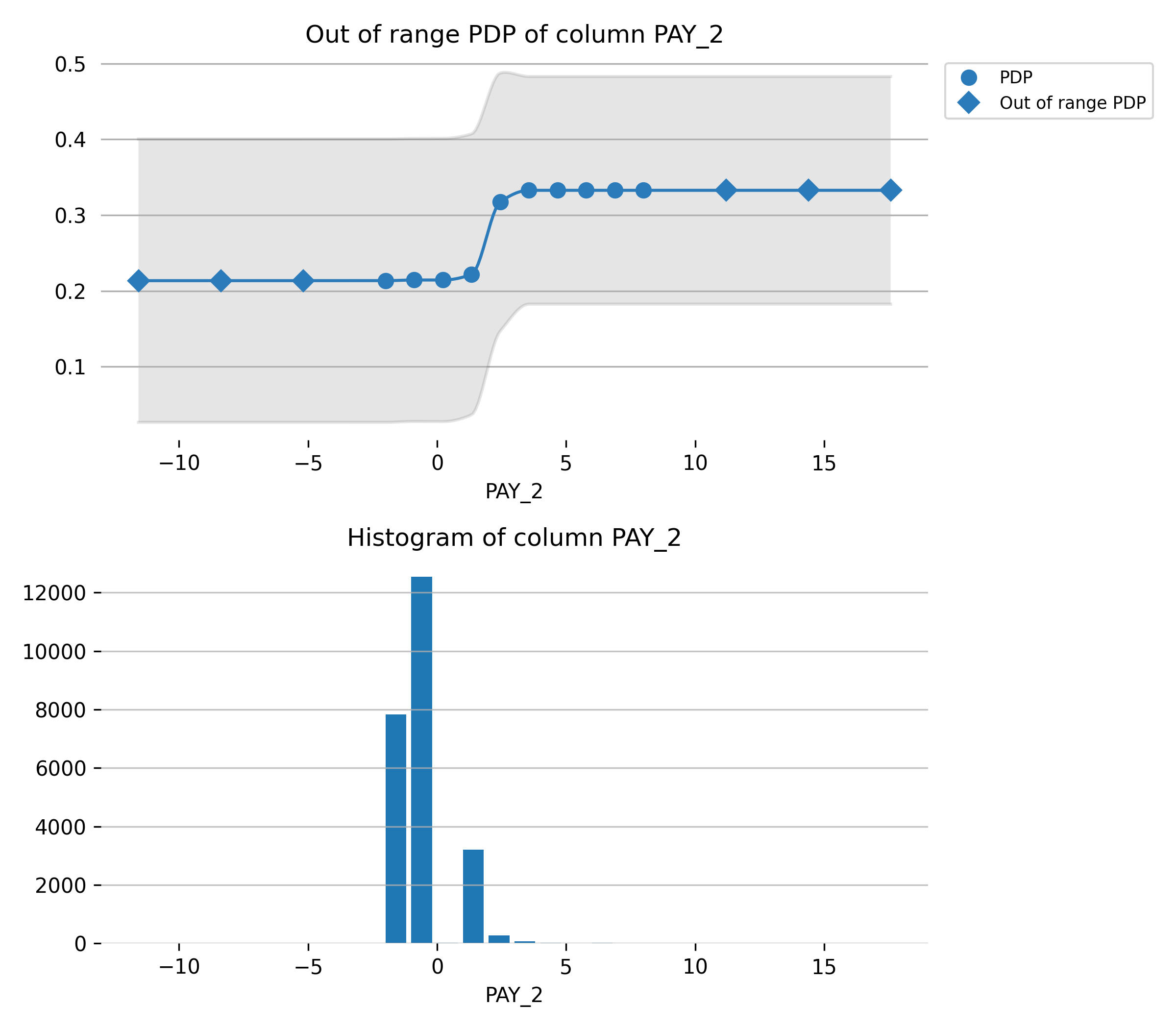
Partial dependence plots show the partial dependence as a function of specific values for a feature subset. The plots show how machine-learned response functions change based on the values of an input feature of interest, while taking nonlinearity into consideration and averaging out the effects of all other input features. Partial dependence plots enable increased transparency in a model and enable the ability to validate and debug a model by comparing a feature’s average predictions across its domain to known standards and reasonable expectations.

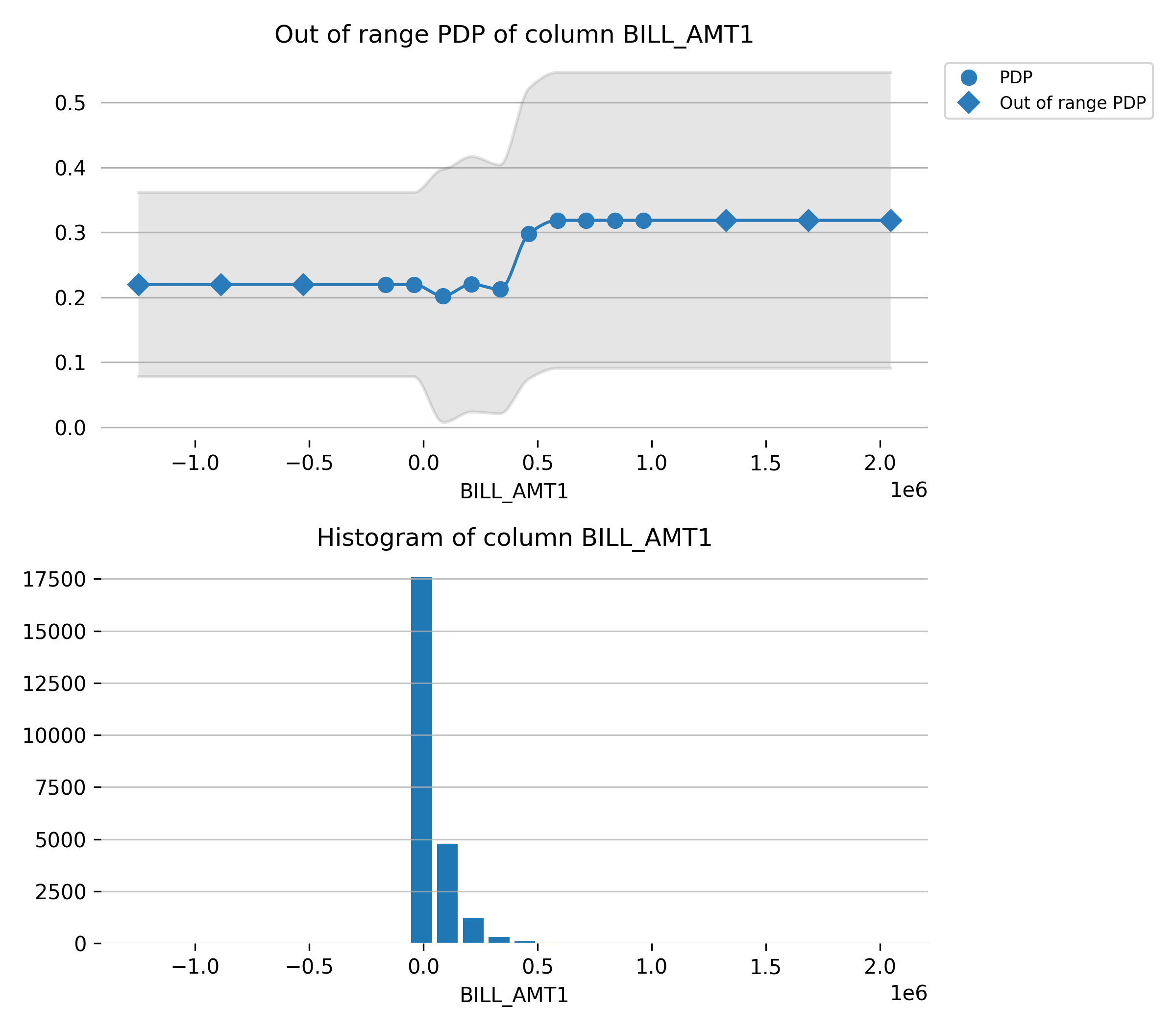
The partial dependence plots are shown for the top 20 original variables. The top 20 original variables are chosen based on their model specific variable importance.

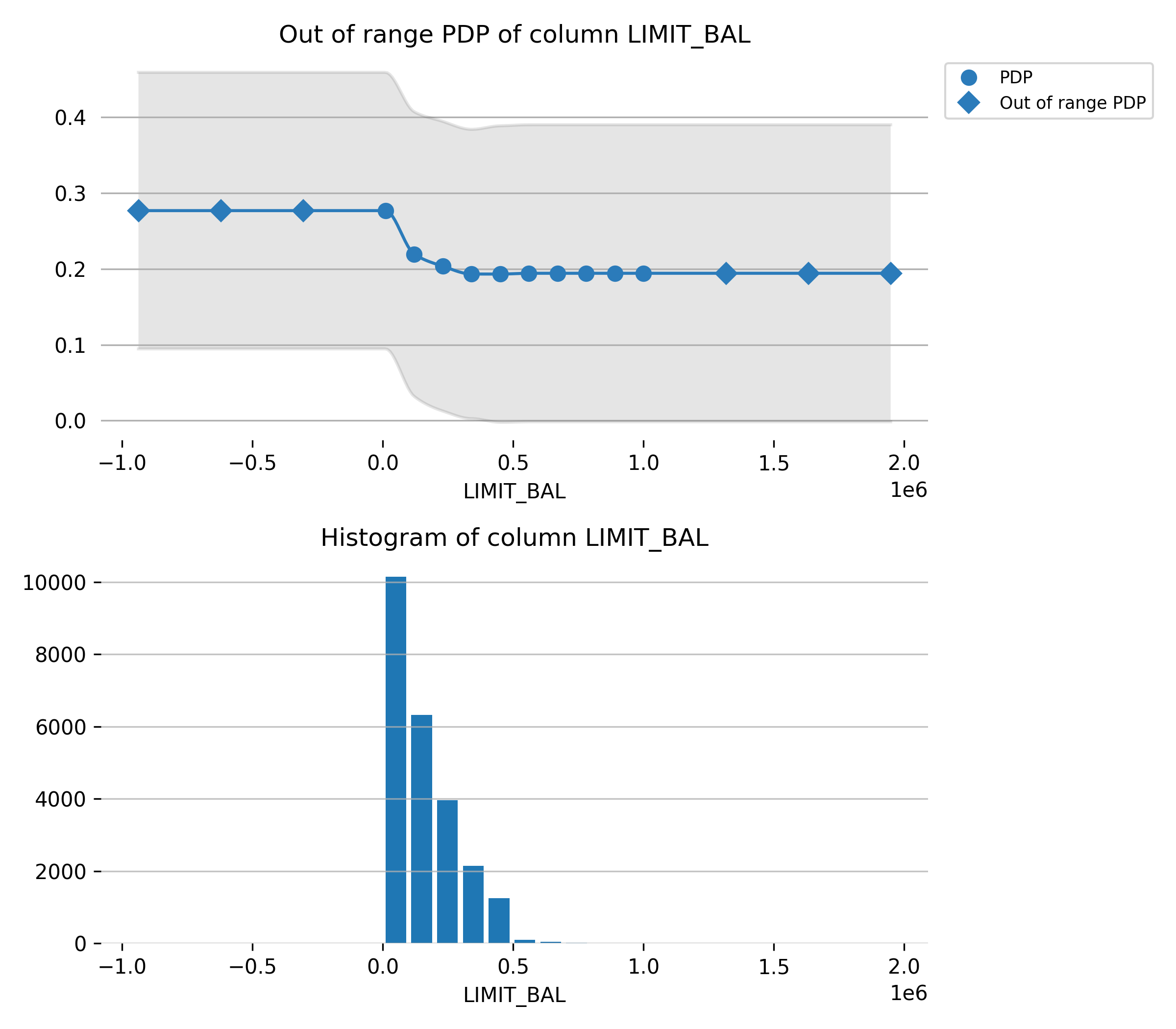
**Plot Details**

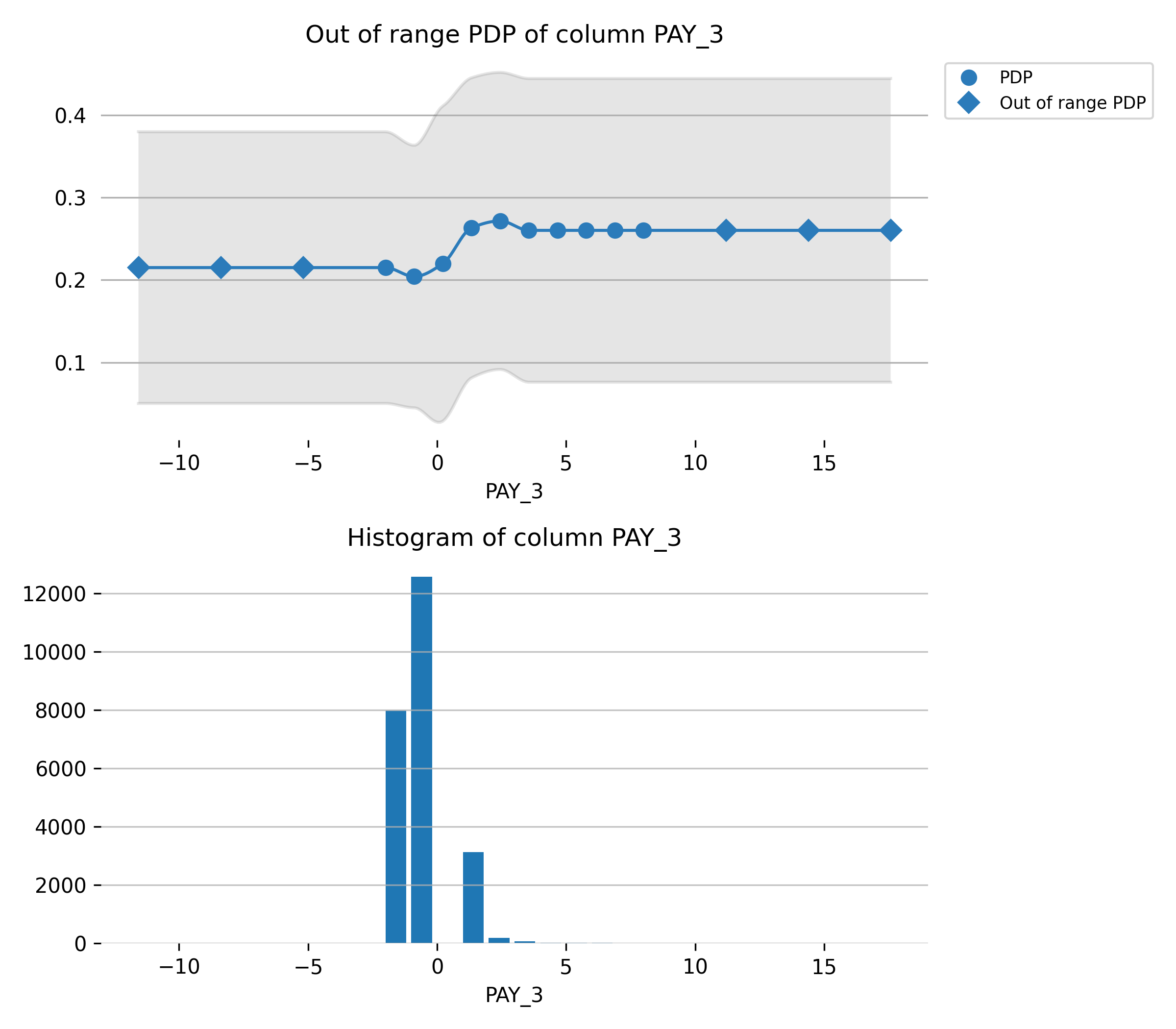
In the H2O-3 PDP, the y-axis represents the mean response, and a shaded region (for numeric features) or shaded bar (for categorical features) represents ± 1 standard deviation.

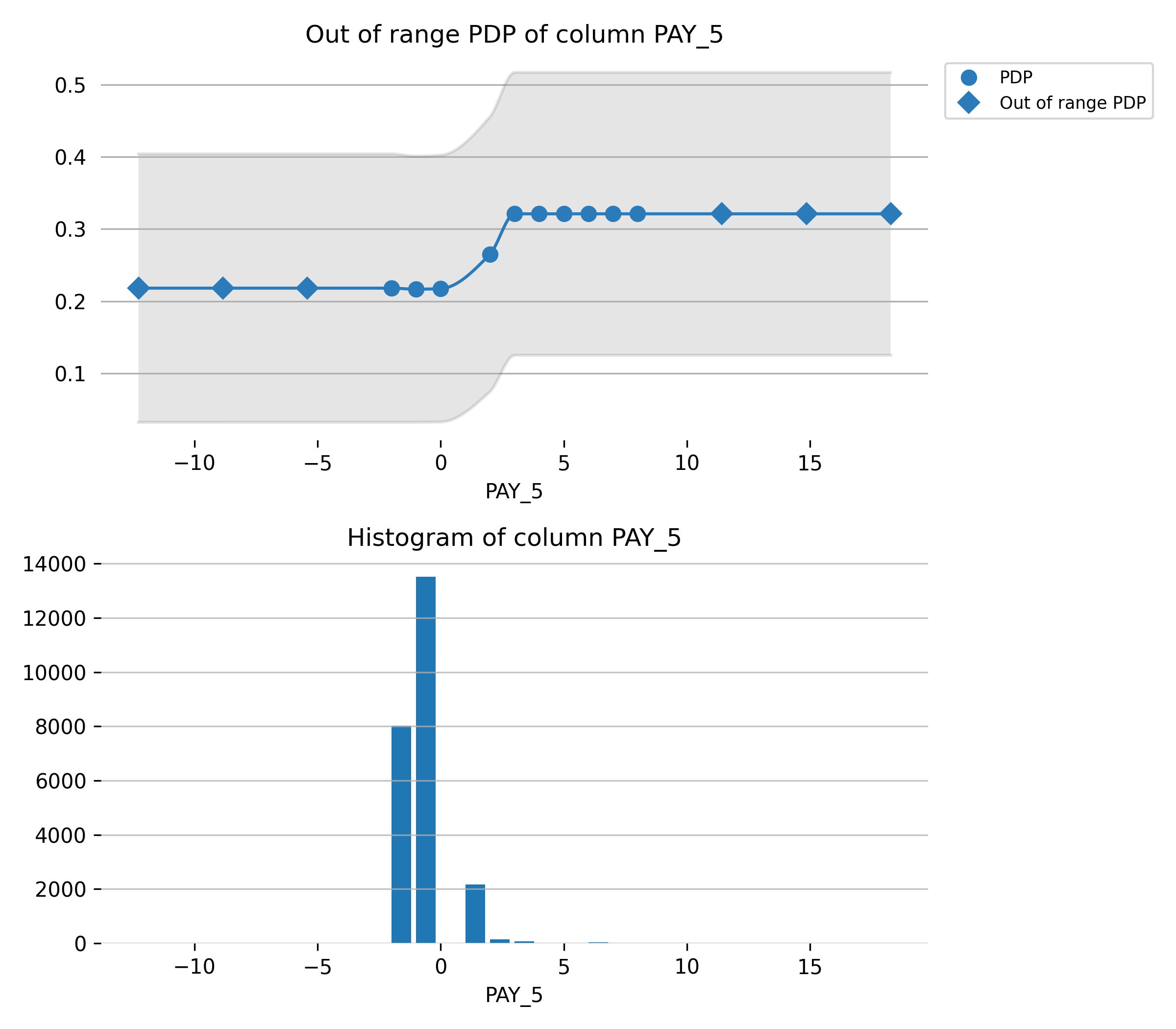
Feature **PAY\_0**

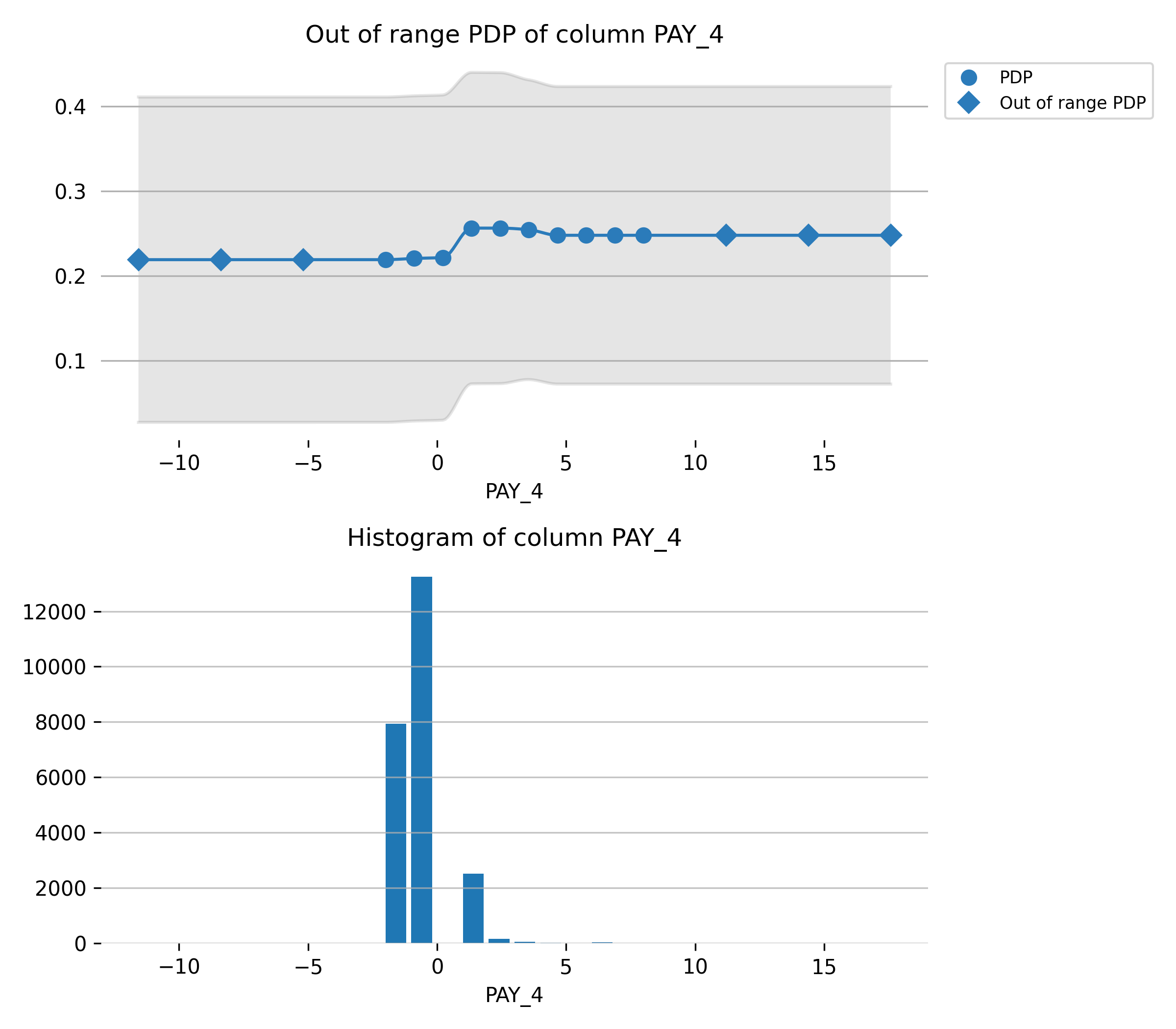
Feature **PAY\_2**

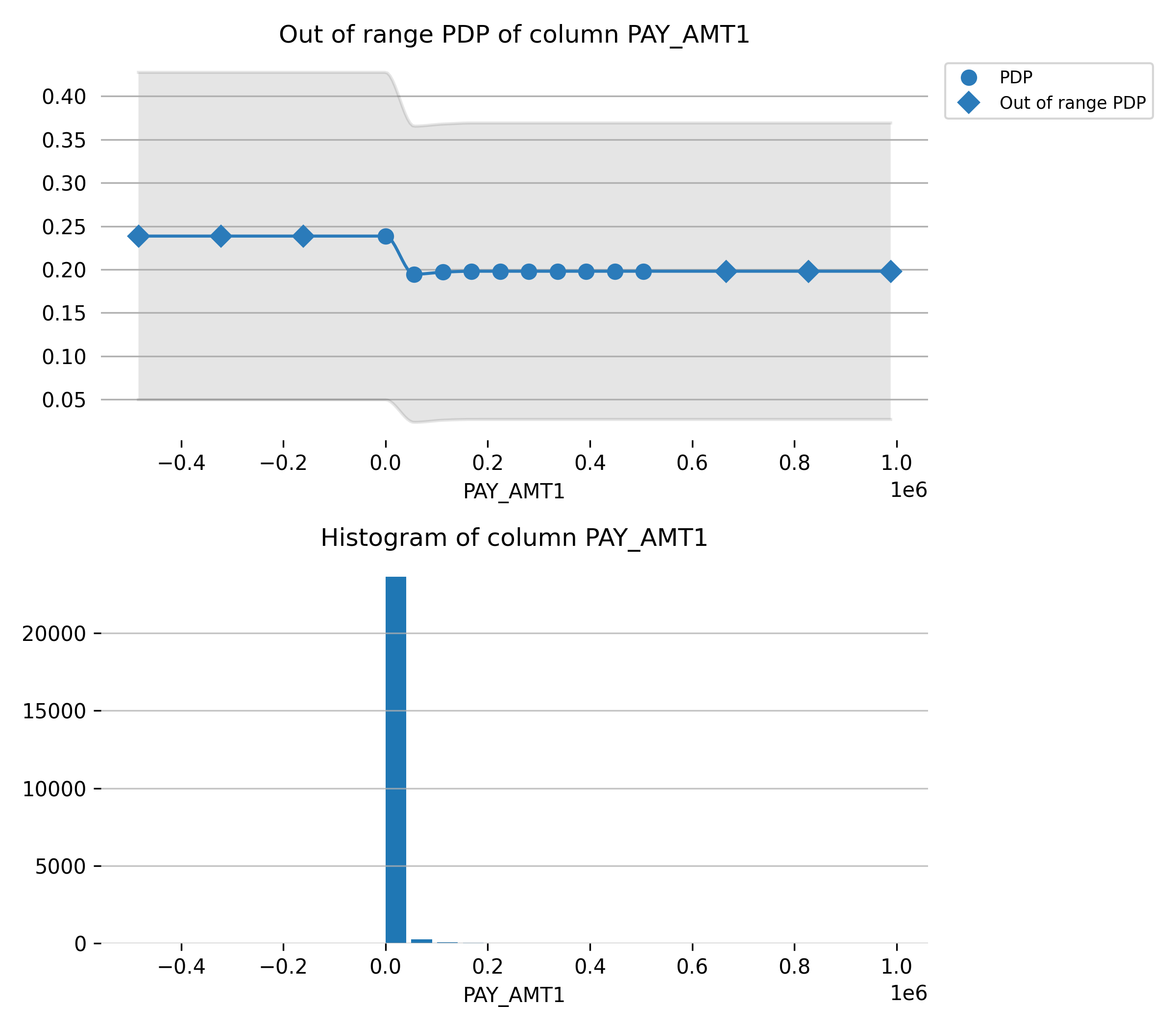
Feature **BILL\_AMT1**

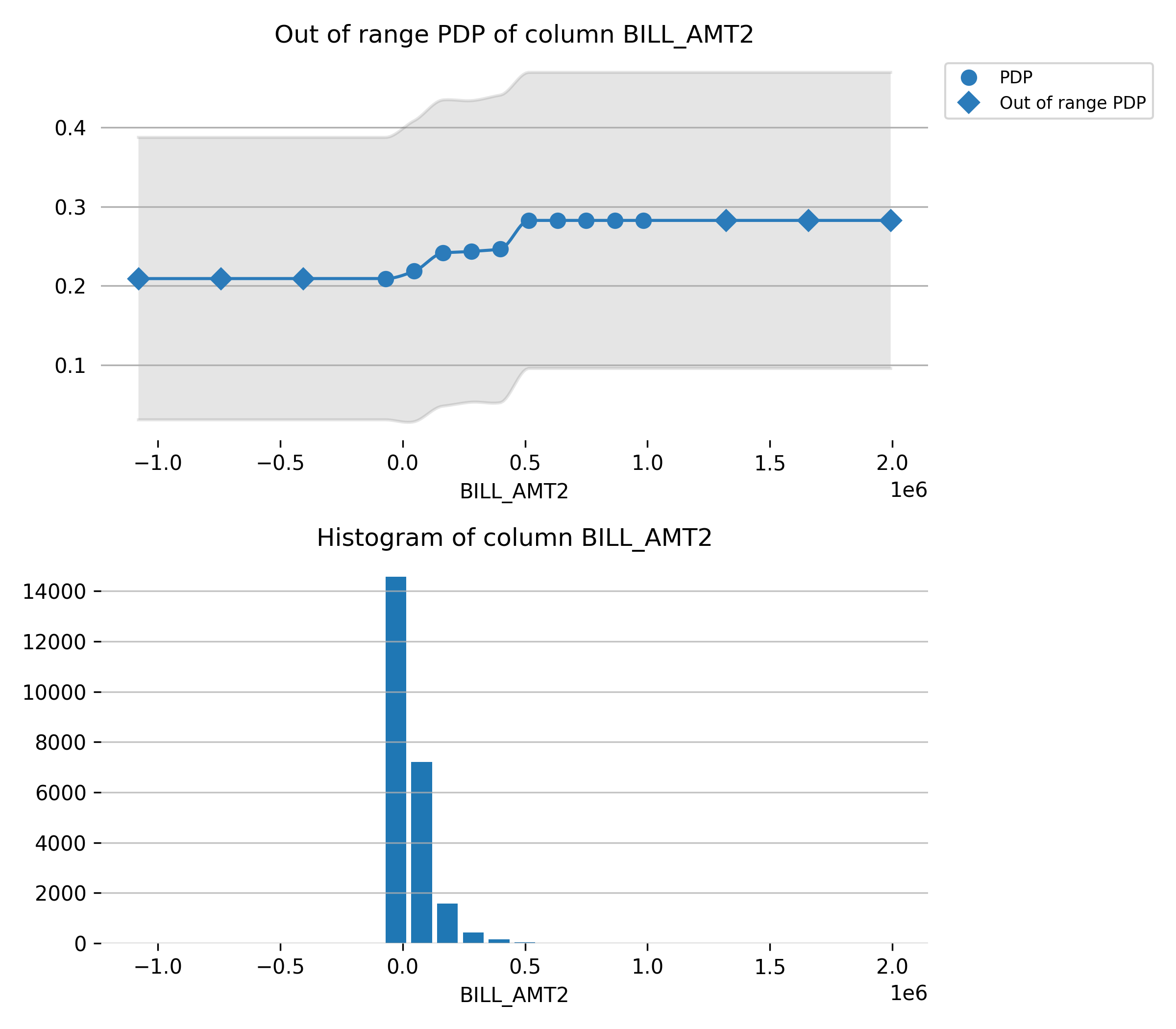
Feature **LIMIT\_BAL**

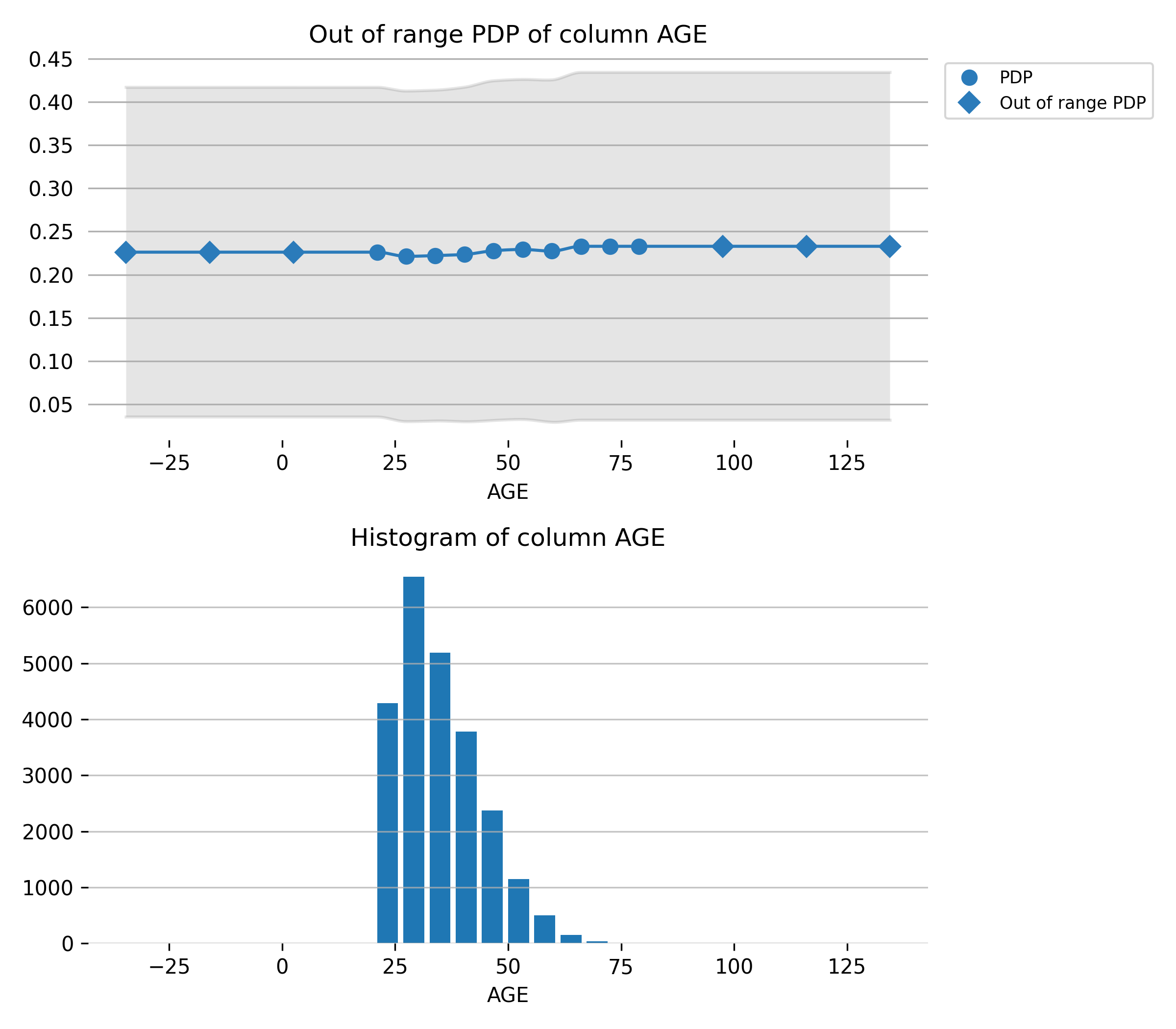
Feature **PAY\_3**

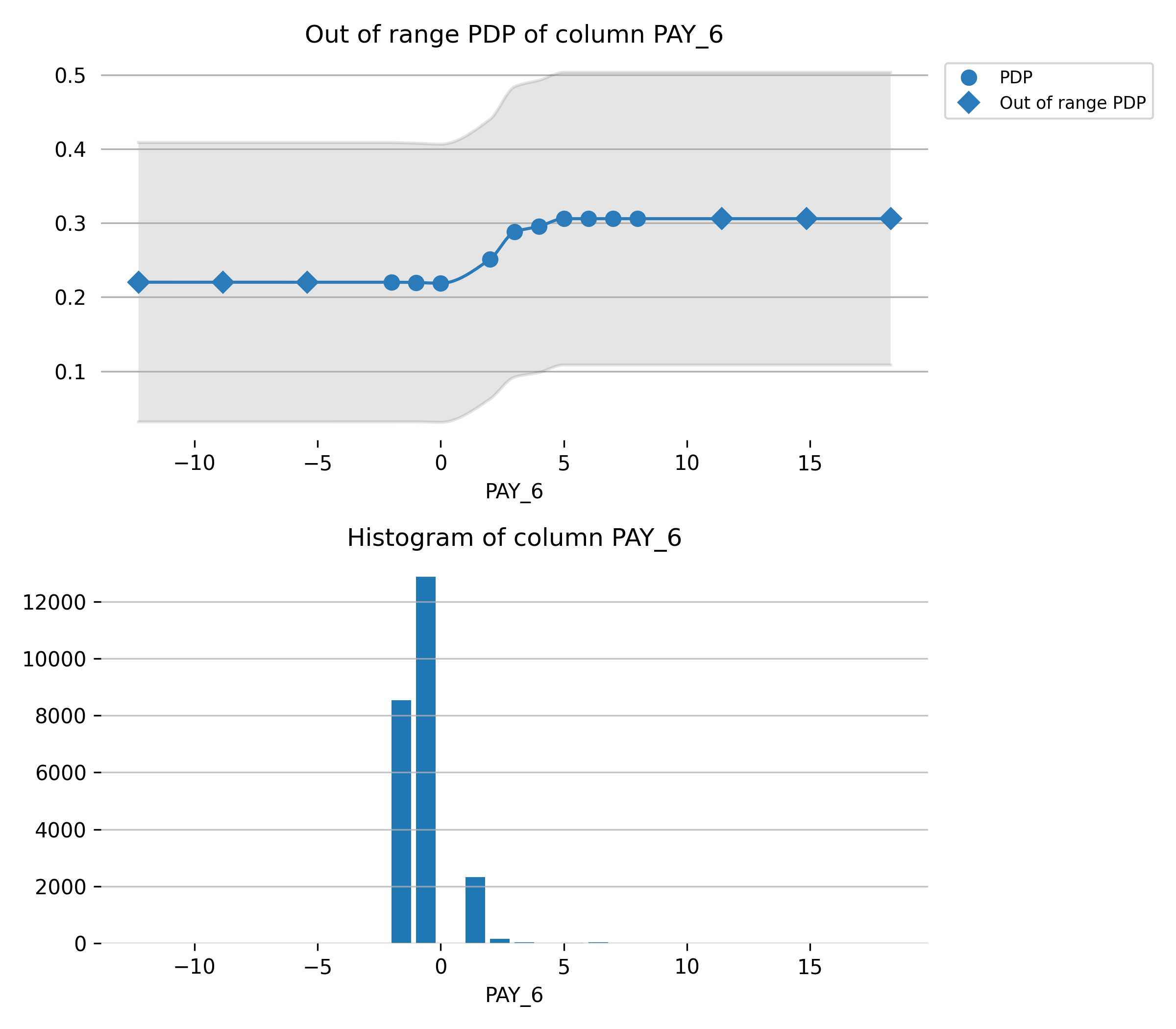
Feature **PAY\_5**

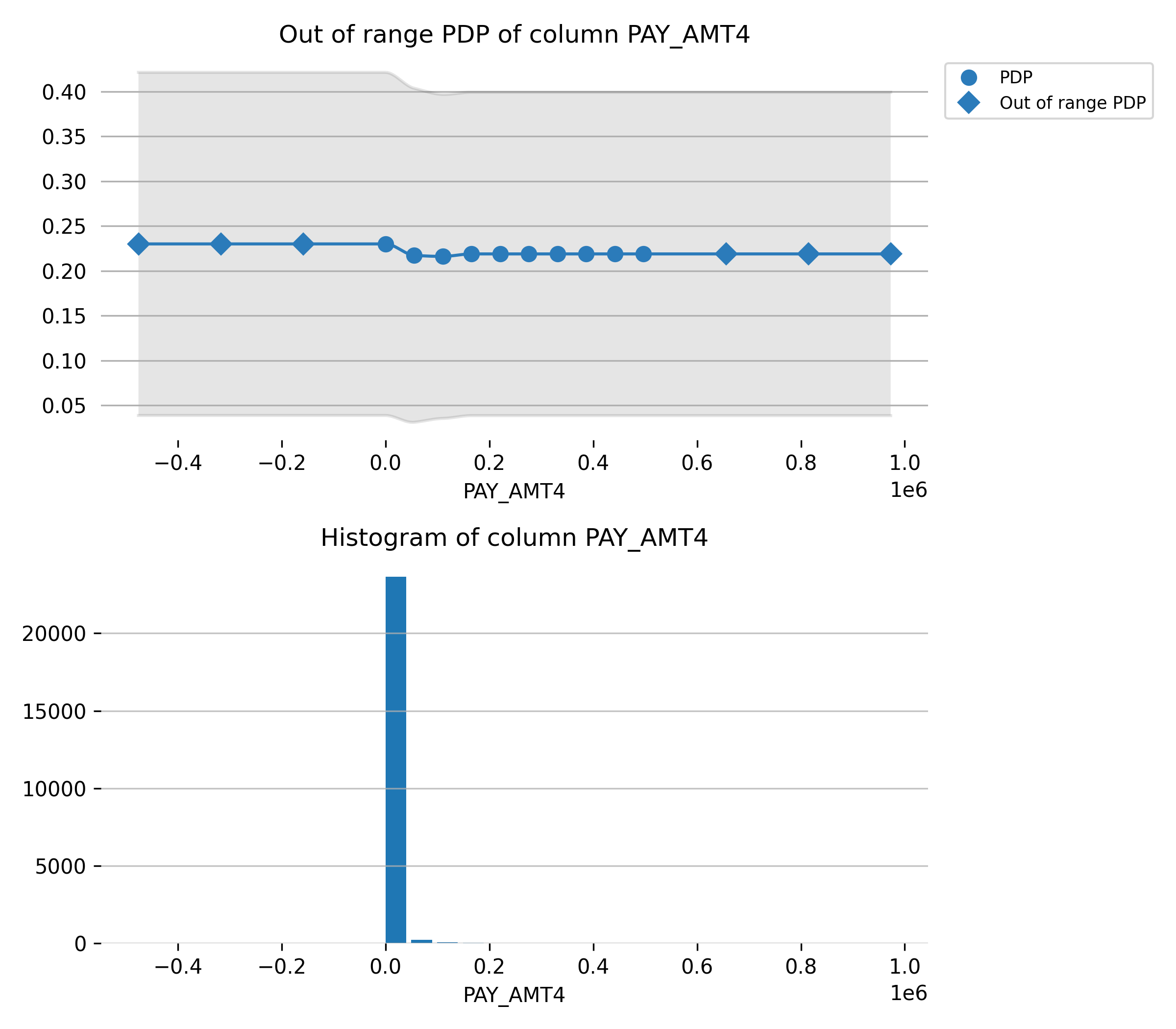
Feature **PAY\_4**

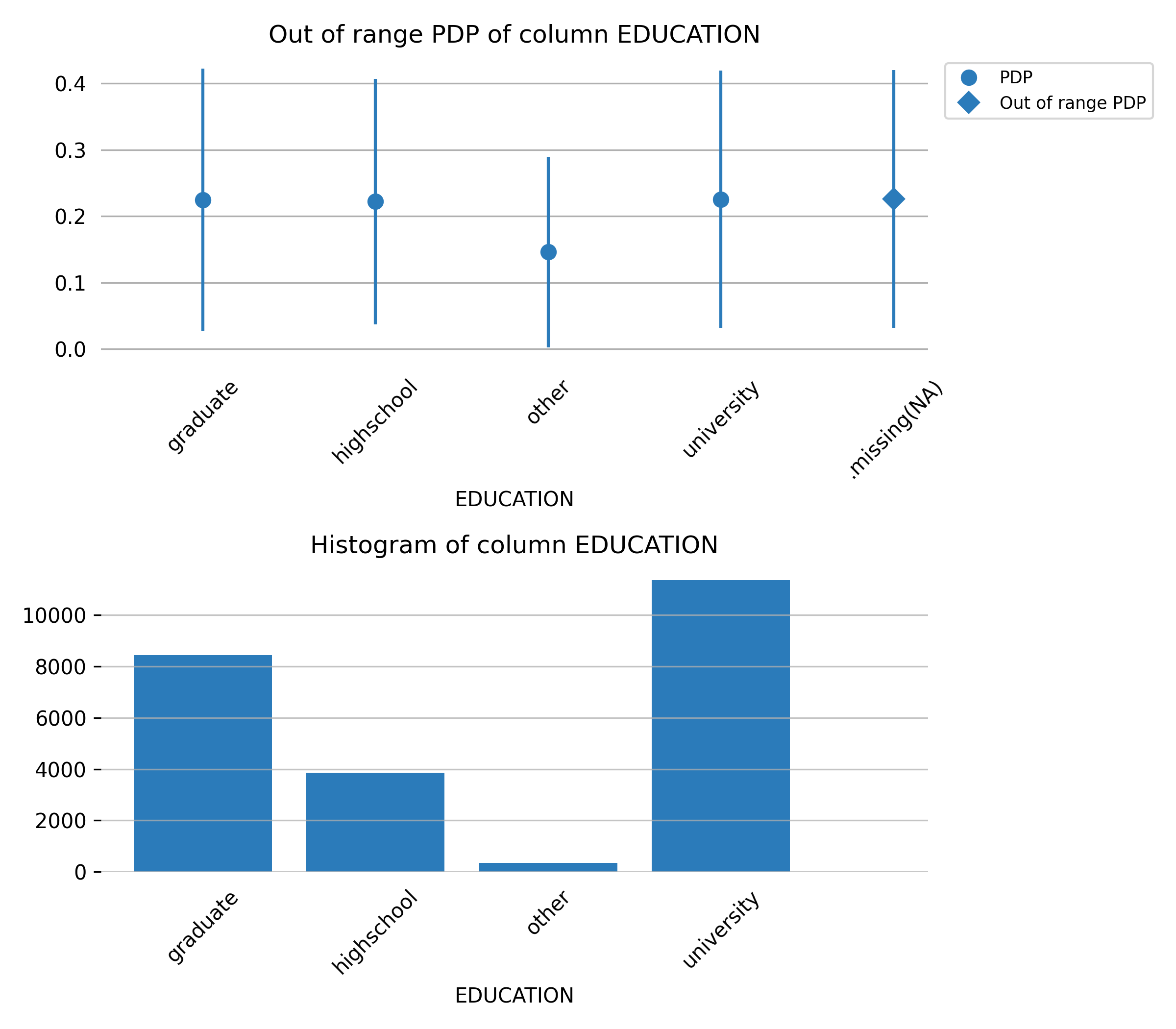
Feature **PAY\_AMT1**

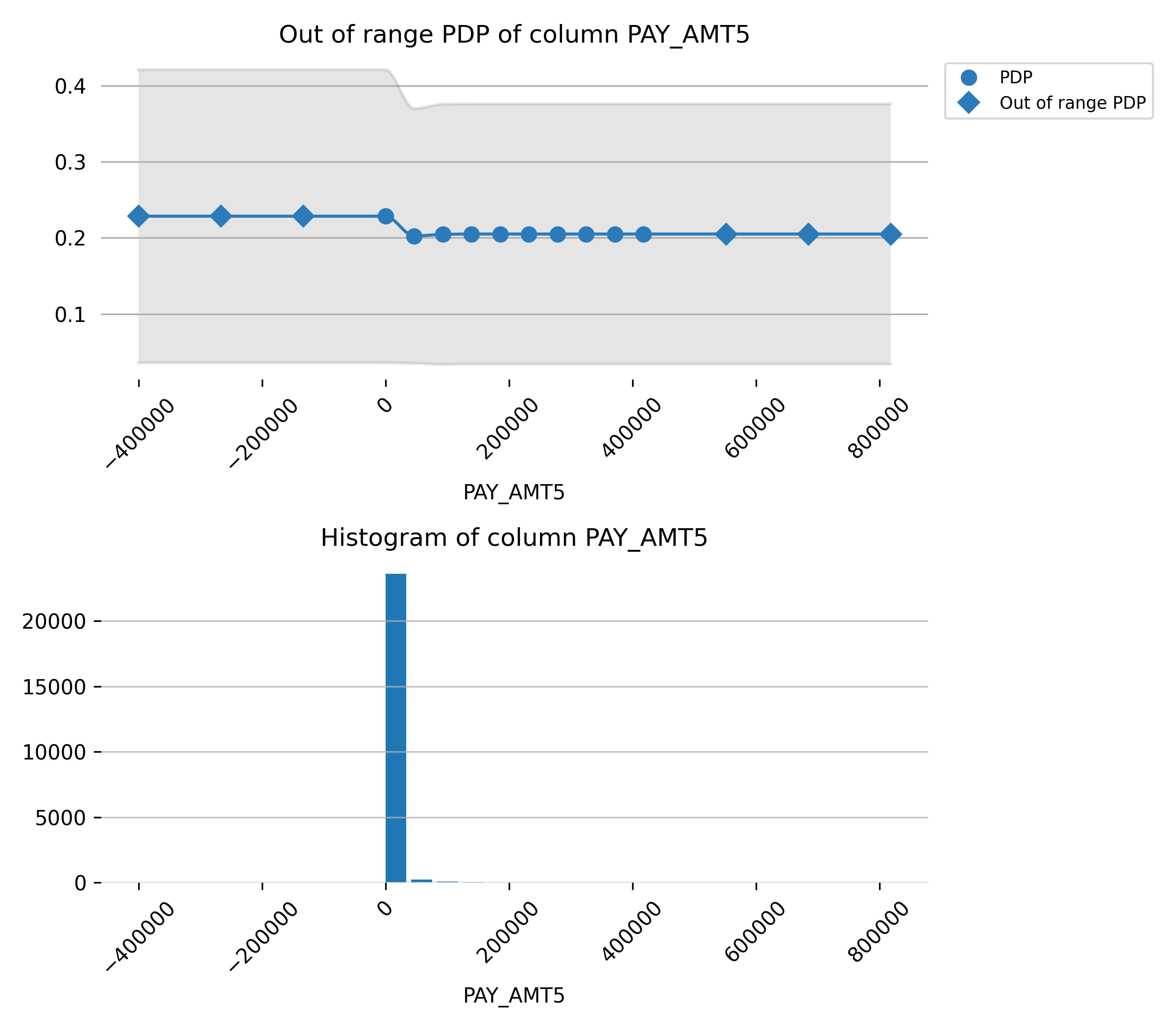
Feature **BILL\_AMT2**

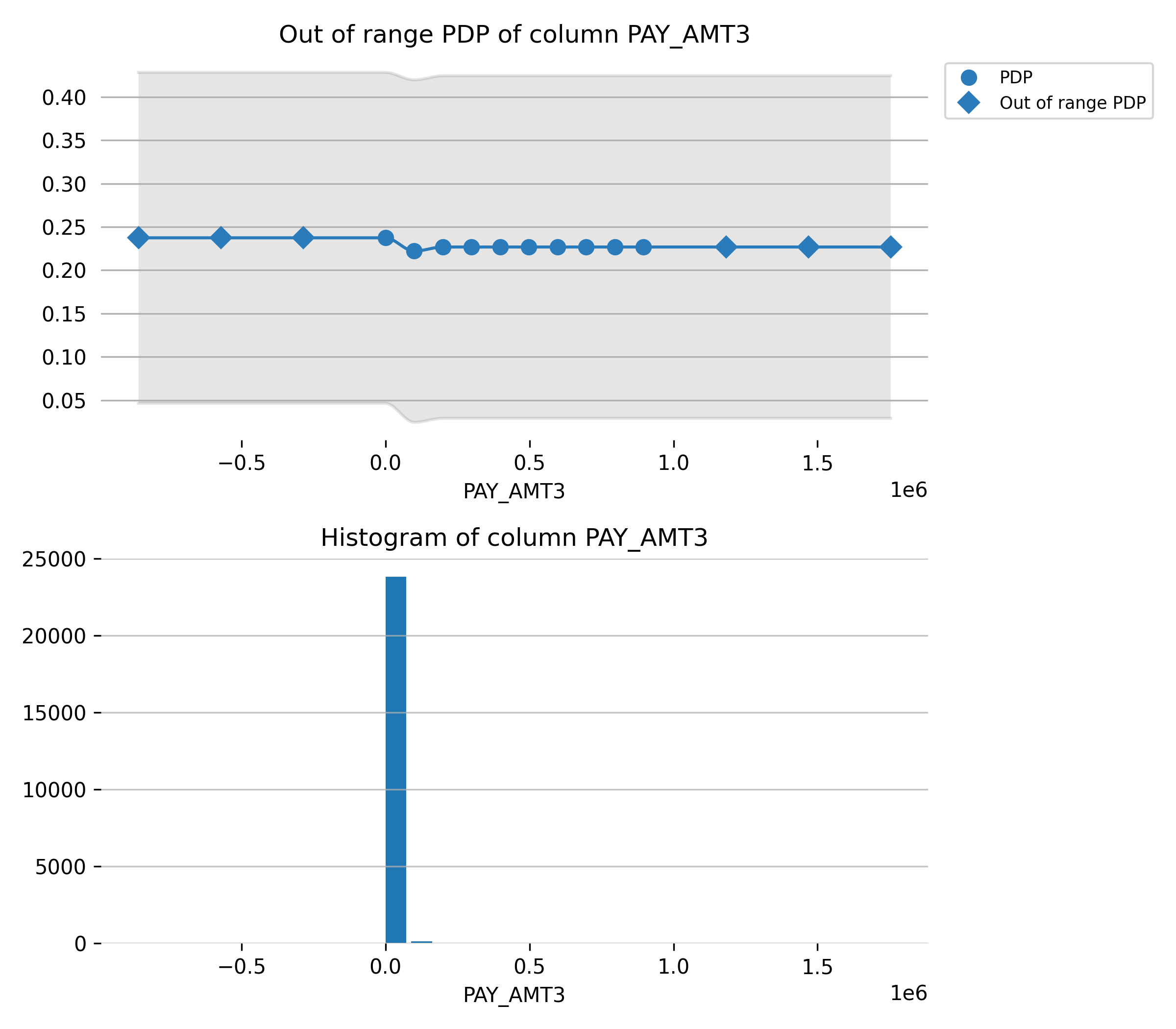
Feature **AGE**

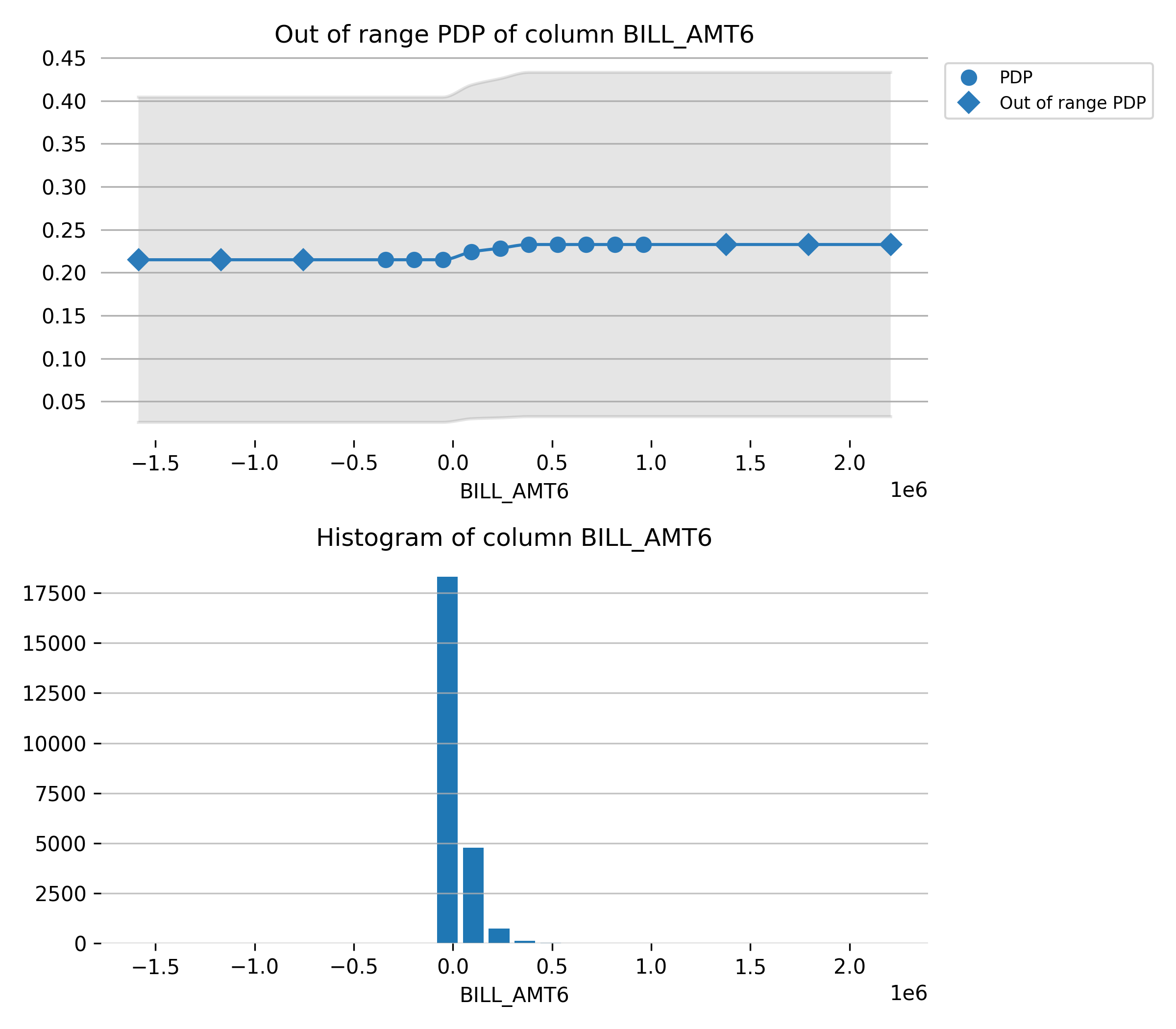
Feature **PAY\_6**

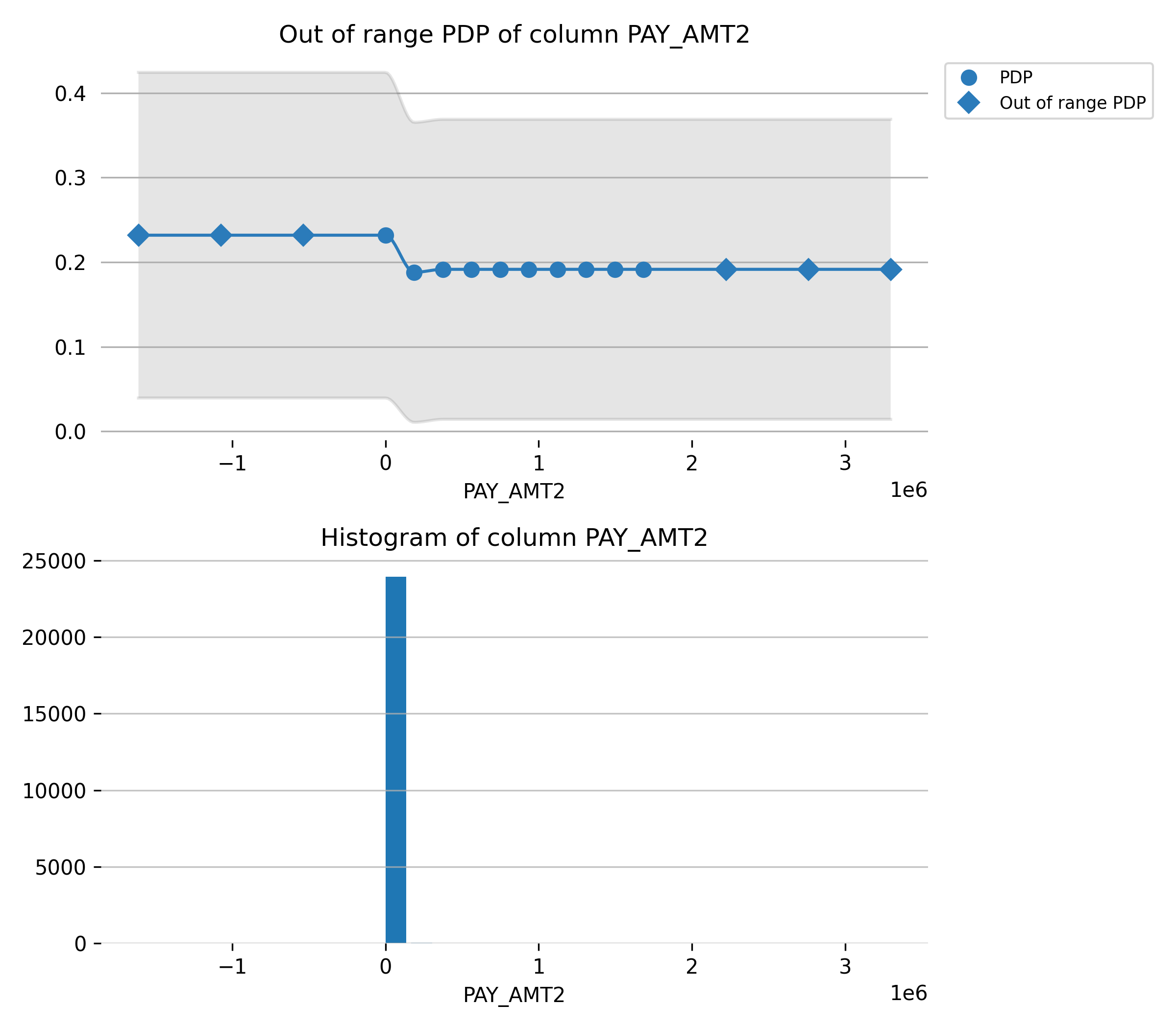
Feature **PAY\_AMT4**

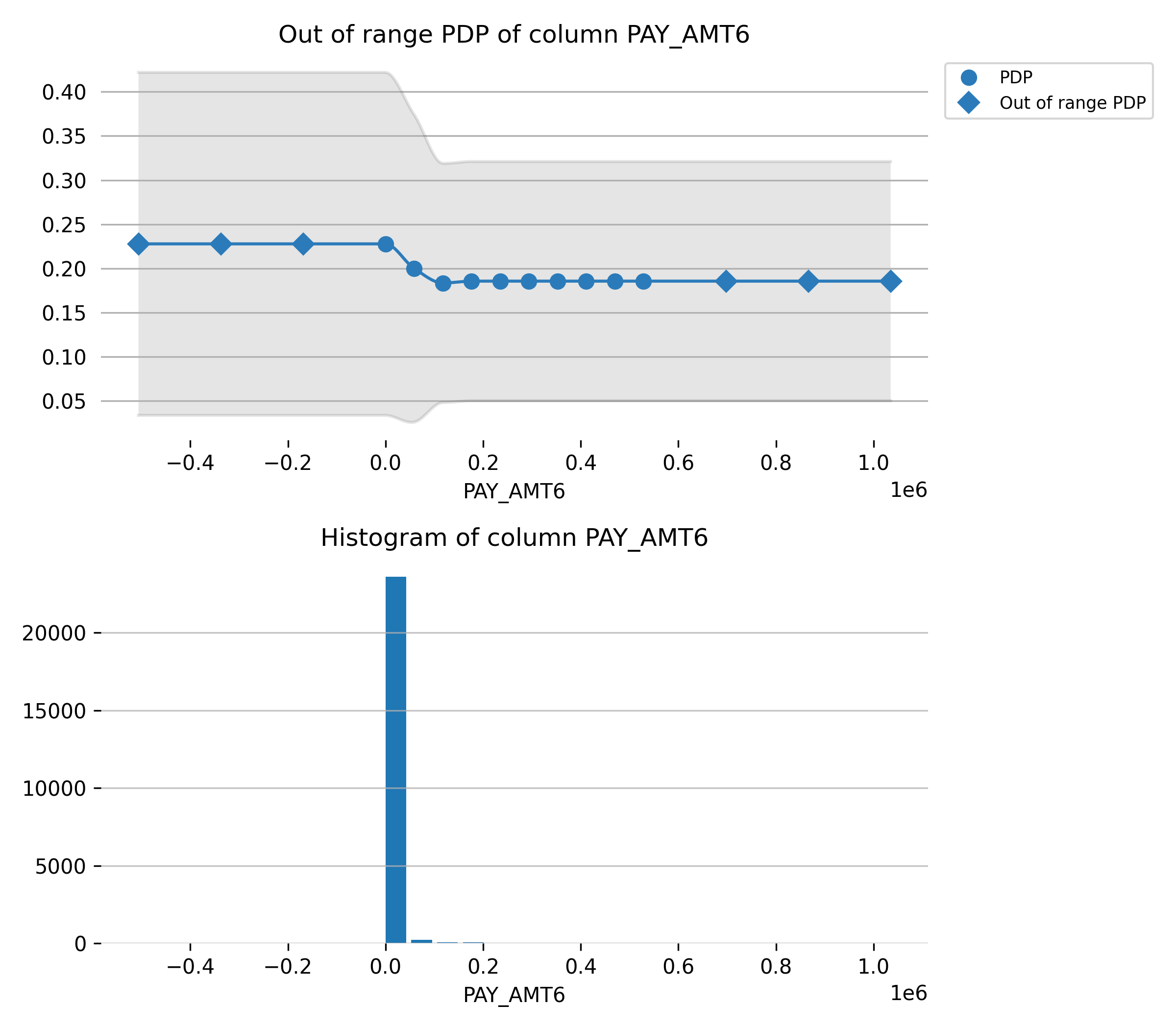
Feature **EDUCATION**

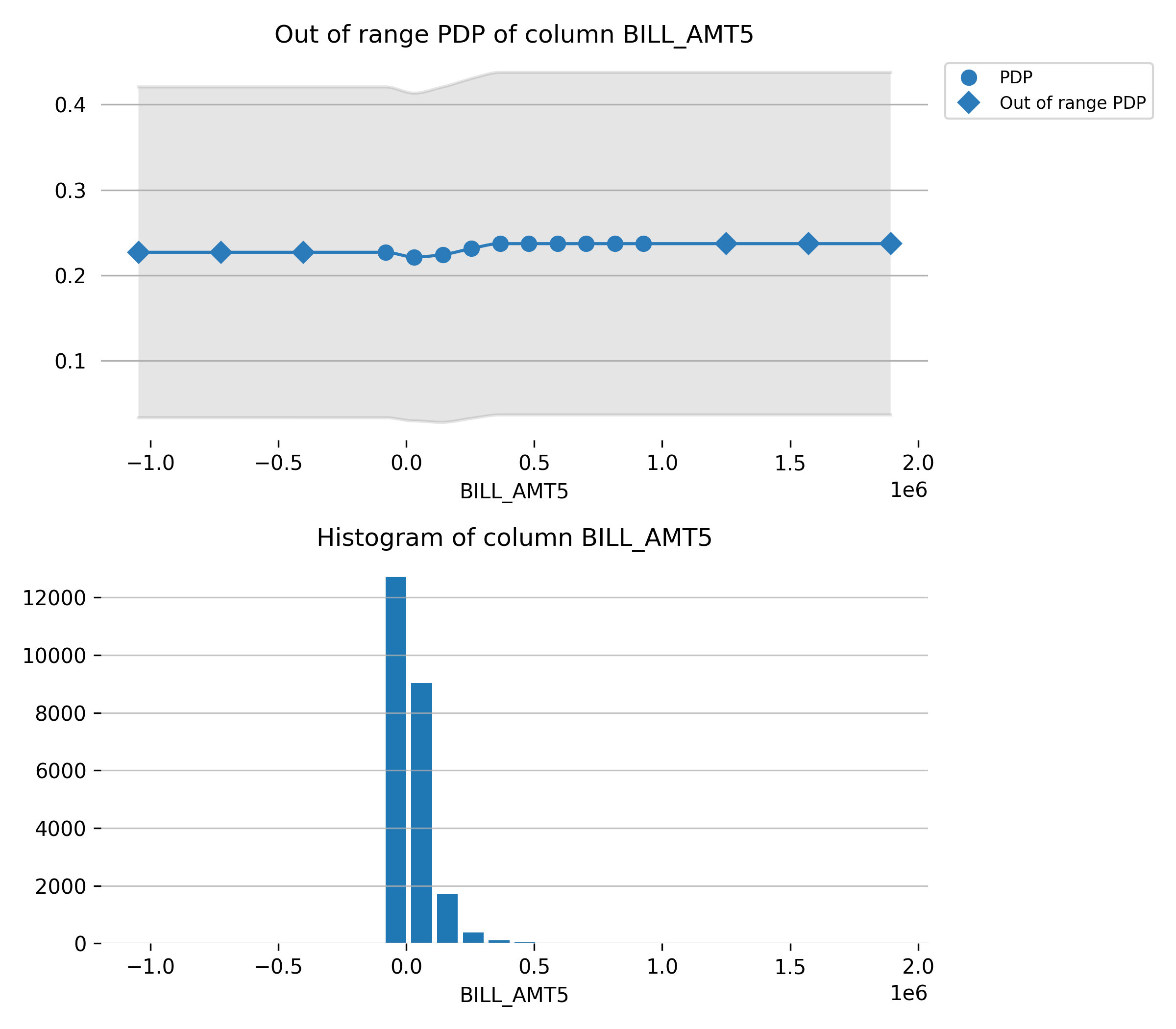
Feature **PAY\_AMT5**

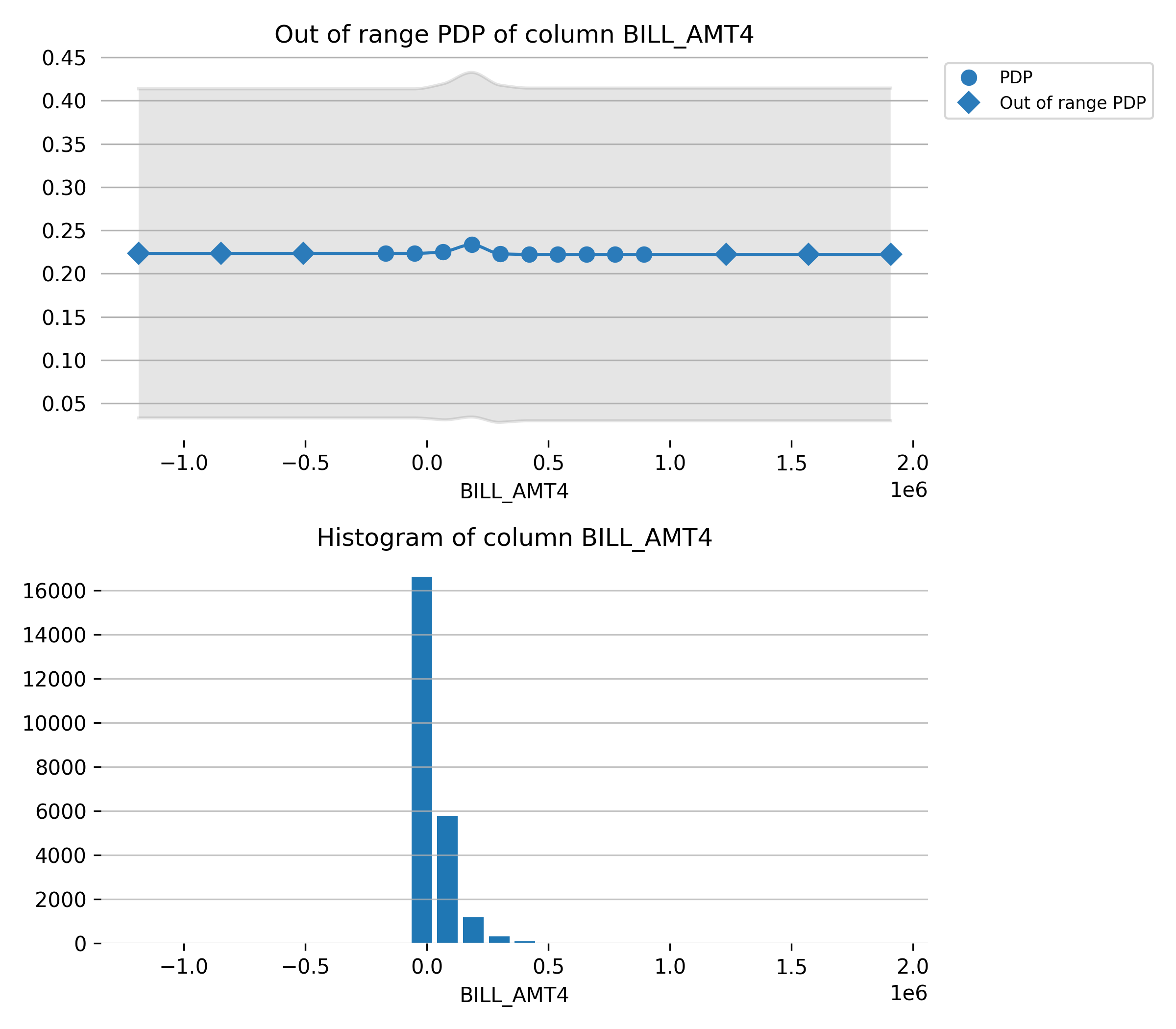
Feature **PAY\_AMT3**

Feature **BILL\_AMT6**

Feature **PAY\_AMT2**

Feature **PAY\_AMT6**

Feature **BILL\_AMT5**

Feature **BILL\_AMT4**

## Model Reproducibility

To reproduce this Gradient Boosting Machine on a single-node cluster, train a model using the same: datasets, model parameters, and cluster configuration.

**Datasets**

Make sure the model is trained on the same datasets. The datasets are the same if it they have the same hash:

* training dataset: -43812237386064560

The *h2o.frame()* function, which retrieves metadata for an H2OFrame’s id, can be used to verify a dataset’s hash:

*h2o.frame(frame.key)[‘frames’][0][‘checksum’]*

**Note:** Reproducibility is only guaranteed for single-file imports. H2O-3 may shuffle the data during a multi-file directory import, and therefore reproducibility cannot be guaranteed.

**Model Parameters**

The same model parameters and H2O-3 version must be used to reproduce this model. Parameters set to ‘AUTO’ use the version’s default values and should not be set by the user.

|  |  |
| --- | --- |
| **Package** | **Version Number** |
| H2O-3 | 3.32.1.7 |

|  |  |
| --- | --- |
| **Model Parameter** | **Value** |
| model\_id | gbm\_model |
| nfolds | 0 |
| keep\_cross\_validation\_models | True |
| keep\_cross\_validation\_predictions | False |
| keep\_cross\_validation\_fold\_assignment | False |
| score\_each\_iteration | False |
| score\_tree\_interval | 0 |
| fold\_assignment | None |
| fold\_column | None |
| response\_column | DEFAULT\_PAYMENT\_NEXT\_MONTH |
| ignored\_columns | ['ID'] |
| ignore\_const\_cols | True |
| offset\_column | None |
| weights\_column | None |
| balance\_classes | False |
| class\_sampling\_factors | None |
| max\_after\_balance\_size | 5.0 |
| max\_confusion\_matrix\_size | 20 |
| ntrees | 50 |
| max\_depth | 5 |
| min\_rows | 10.0 |
| nbins | 20 |
| nbins\_top\_level | 1024 |
| nbins\_cats | 1024 |
| r2\_stopping | 1.7976931348623157e+308 |
| stopping\_rounds | 0 |
| stopping\_metric | None |
| stopping\_tolerance | 0.001 |
| max\_runtime\_secs | 0.0 |
| seed | 1234 |
| build\_tree\_one\_node | False |
| learn\_rate | 0.1 |
| learn\_rate\_annealing | 1.0 |
| distribution | bernoulli |
| quantile\_alpha | 0.5 |
| tweedie\_power | 1.5 |
| huber\_alpha | 0.9 |
| checkpoint | None |
| sample\_rate | 1.0 |
| sample\_rate\_per\_class | None |
| col\_sample\_rate | 1.0 |
| col\_sample\_rate\_change\_per\_level | 1.0 |
| col\_sample\_rate\_per\_tree | 1.0 |
| min\_split\_improvement | 1e-05 |
| histogram\_type | UniformAdaptive |
| max\_abs\_leafnode\_pred | 1.7976931348623157e+308 |
| pred\_noise\_bandwidth | 0.0 |
| categorical\_encoding | Enum |
| calibrate\_model | False |
| calibration\_frame | None |
| custom\_metric\_func | None |
| custom\_distribution\_func | None |
| export\_checkpoints\_dir | None |
| monotone\_constraints | None |
| check\_constant\_response | True |
| gainslift\_bins | -1 |
| auc\_type | AUTO |

**Cluster Configuration**

The same cluster configuration is required.

|  |  |
| --- | --- |
| **Cluster Parameters** | **Value** |
| Total Nodes | 1 |
| Total Allowed Cores | 8 |

**Node Order**

The cluster’s leader node must trigger the model training.

|  |  |  |
| --- | --- | --- |
| **Node Index** | **Cluster Leader** | **Node IP** |
| 0 | True | 127.0.0.1:54321 |

**Note:** When H2O-3 is running on Hadoop, the h2odriver automatically returns the leader node to which the user should connect. In multi-node deployments of Standalone H2O-3, the user must identify the leader node. Flow users can easily check whether they are connected to the leader node by opening Cluster Status (from the Admin menu) and checking that the first node has the same IP address as they see in their browser’s address bar.

## Appendix

### Final Model Details

|  |  |
| --- | --- |
| **Model Parameters (Complete List)** | **Values** |
| model\_id | gbm\_model |
| auc\_type | AUTO |
| balance\_classes | False |
| build\_tree\_one\_node | False |
| calibrate\_model | False |
| calibration\_frame | None |
| categorical\_encoding | Enum |
| check\_constant\_response | True |
| checkpoint | None |
| class\_sampling\_factors | None |
| col\_sample\_rate | 1.0 |
| col\_sample\_rate\_change\_per\_level | 1.0 |
| col\_sample\_rate\_per\_tree | 1.0 |
| custom\_distribution\_func | None |
| custom\_metric\_func | None |
| distribution | bernoulli |
| export\_checkpoints\_dir | None |
| fold\_assignment | None |
| fold\_column | None |
| gainslift\_bins | -1 |
| histogram\_type | UniformAdaptive |
| ignore\_const\_cols | True |
| ignored\_columns | ['ID'] |
| keep\_cross\_validation\_fold\_assignment | False |
| keep\_cross\_validation\_models | True |
| keep\_cross\_validation\_predictions | False |
| learn\_rate | 0.1 |
| learn\_rate\_annealing | 1.0 |
| max\_abs\_leafnode\_pred | 1.7976931348623157e+308 |
| max\_after\_balance\_size | 5.0 |
| max\_confusion\_matrix\_size | 20 |
| max\_depth | 5 |
| max\_runtime\_secs | 0.0 |
| min\_rows | 10.0 |
| min\_split\_improvement | 1e-05 |
| monotone\_constraints | None |
| nbins | 20 |
| nbins\_cats | 1024 |
| nbins\_top\_level | 1024 |
| nfolds | 0 |
| ntrees | 50 |
| offset\_column | None |
| pred\_noise\_bandwidth | 0.0 |
| response\_column | DEFAULT\_PAYMENT\_NEXT\_MONTH |
| sample\_rate | 1.0 |
| sample\_rate\_per\_class | None |
| score\_each\_iteration | False |
| score\_tree\_interval | 0 |
| seed | 1234 |
| stopping\_metric | None |
| stopping\_rounds | 0 |
| stopping\_tolerance | 0.001 |
| weights\_column | None |

**Population Stability Index (PSI) Final Model Details**

Population Stability Index is a statistic used to describe a variable’s distribution shift. It can measure the shift between the training dataset’s model score distribution and any other given dataset (i.e. validation or test dataset).

A PSI value lower than 0.10 indicates a small shift in the model predictions, a value between 0.10 and 0.25 indicates a moderate shift, and a value greater than 0.25 indicates a strong shift. Strong shift values can indicate that the model trained on the training dataset might not be suitable for the provided validation or test datasets.

The PSI and calculation table is provided for each dataset below. The corresponding table columns are defined as follows:

* *Quantile: the bin to which the ordered predicted probabilities belong.*
* *Upper Bound: the upper bound of the corresponding bin.*
* *Test Count: the total number of Test records within the corresponding bin.*
* *Test Fraction (Tst): Test Count divided by the total number of Test records.*
* *Train Count: the total number of Train records within the corresponding bin.*
* *Train Fraction (Trn): Train Count divided by the total number of Train records.*
* *Tst - Trn: the difference between the Test Fraction and the Train Fraction.*
* *ln(Tst / Trn): the natural logarithm of the Test Fraction divided by the Train Fraction.*
* *PSI: the Population Stability Index for each bin - the dataset PSI is the total sum of these PSI values.*

**Validation**

The Population Stability Index is 0.005171.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quantile** | **Upper Bound** | **Test Count** | **Test Fraction (Tst)** | **Train Count** | **Train Fraction (Trn)** | **Tst - Trn** | **ln(Tst / Trn)** | **PSI** |
| 0.1 | 0.0721 | 612 | 0.102 | 2,399 | 0.1 | 0.002 | 0.0202 | 0.0 |
| 0.2 | 0.0902 | 628 | 0.1047 | 2,400 | 0.1 | 0.0047 | 0.0456 | 0.0002 |
| 0.3 | 0.1066 | 667 | 0.1112 | 2,401 | 0.1 | 0.0111 | 0.1054 | 0.0012 |
| 0.4 | 0.1264 | 646 | 0.1077 | 2,400 | 0.1 | 0.0077 | 0.0738 | 0.0006 |
| 0.5 | 0.1475 | 592 | 0.0987 | 2,400 | 0.1 | -0.0013 | -0.0135 | 0.0 |
| 0.6 | 0.1745 | 597 | 0.0995 | 2,399 | 0.1 | -0.0005 | -0.0046 | 0.0 |
| 0.7 | 0.2204 | 629 | 0.1048 | 2,400 | 0.1 | 0.0048 | 0.0472 | 0.0002 |
| 0.8 | 0.3169 | 553 | 0.0922 | 2,392 | 0.0997 | -0.0075 | -0.0783 | 0.0006 |
| 0.9 | 0.5891 | 544 | 0.0907 | 2,408 | 0.1003 | -0.0097 | -0.1013 | 0.001 |
| 1.0 | inf | 532 | 0.0887 | 2,400 | 0.1 | -0.0113 | -0.1203 | 0.0014 |

**Quantile Plots Calculation Table**

The following table is used to calculate the Quantile Response Rates, Actual vs Predicted Probabilities, and Actual vs Predicted Log Odds plot. table columns are defined as follows:

* *Quantile: the bin to which the ordered predicted probabilities belong.*
* *bound: the upper bound of the corresponding bin.*
* *{dataset name} cnt: the number of records within the corresponding bin.*
* *{dataset name} sum: the number of positive-labeled records within the corresponding bin.*
* *{dataset name} act: the fraction of positive-labeled records within the corresponding bin.*
* *{dataset name} pred: the mean of the predicted values that fall within the corresponding bin.*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Quantile** | **bound** | **Train cnt** | **Train sum** | **Train act** | **Train pred** | **Validation cnt** | **Validation sum** | **Validation act** | **Validation pred** |
| 0.1 | inf | 2,400 | 1,834 | 0.7642 | 0.7079 | 532 | 388 | 0.7293 | 0.7054 |
| 0.2 | 0.5891 | 2,408 | 1,056 | 0.4385 | 0.4212 | 544 | 246 | 0.4522 | 0.4165 |
| 0.3 | 0.3169 | 2,392 | 683 | 0.2855 | 0.2606 | 553 | 146 | 0.264 | 0.2613 |
| 0.4 | 0.2204 | 2,400 | 446 | 0.1858 | 0.1965 | 629 | 127 | 0.2019 | 0.1961 |
| 0.5 | 0.1745 | 2,399 | 406 | 0.1692 | 0.1598 | 597 | 111 | 0.1859 | 0.1598 |
| 0.6 | 0.1475 | 2,400 | 322 | 0.1342 | 0.1368 | 592 | 82 | 0.1385 | 0.1363 |
| 0.7 | 0.1264 | 2,400 | 246 | 0.1025 | 0.1162 | 646 | 59 | 0.0913 | 0.1165 |
| 0.8 | 0.1066 | 2,401 | 200 | 0.0833 | 0.0983 | 667 | 67 | 0.1004 | 0.0985 |
| 0.9 | 0.0902 | 2,400 | 111 | 0.0462 | 0.0814 | 628 | 22 | 0.035 | 0.0813 |
| 1.0 | 0.0721 | 2,399 | 65 | 0.0271 | 0.0592 | 612 | 18 | 0.0294 | 0.0588 |