Getting to Know SAS® Viya

VA, VS, VDDML and VTA
Agenda

• Overview of Viya
• Introduction and Demo
  • SAS Visual Analytics
  • SAS Visual Statistics
  • SAS Visual Data Mining and Machine Learning
  • SAS Visual Text Analytics
Our digital transformation to power the analytics economy
What is SAS Viya?

SAS Viya is a cloud-enabled, in-memory analytics engine that provides quick, accurate and reliable analytical insights.
SAS Viya Products

- SAS Viya is an underlying foundation for additional products that will take advantage of a cloud-enabled, open platform. Most offerings include both a coding interface as well a visual interface.
  - SAS Visual Analytics
  - SAS Visual Statistics
  - SAS Visual Data Mining and Machine Learning
  - SAS Visual Forecasting
  - SAS Visual Text Mining
  - SAS Optimization
  - SAS Econometrics
  - SAS Visual Investigator
How is SAS Viya Accessed?
MULTIPLE INTERFACES, SINGLE CODE BASE

Visual Interfaces

Programming Interfaces

API Interfaces
Introduction: Software Overview

Multiple Interfaces Target Different Users
What is CAS?
Cloud Analytics Services
Cloud Analytic Services (CAS) In-Memory Engine

The CAS in-memory engine is a fast, scalable, and resilient run-time environment for data management and analytics for SAS® Viya™.

**Fast**
- Multi-threaded
- Distributed In-Memory
- Efficient Inter-node Communication

**Scalable**
- Single Machine to Distributed MPP
- Memory-mapping to process data larger than physical memory * (No memory failures)

**Resilient**
- Fault-tolerant to Node and Network Failures
- Worker and Controller* Failover
- Session Independence
Multiple interfaces, single code base

Clients ask CAS to run “actions” on data

Visual Interfaces

Programming Interfaces

API Interfaces

CAS Server

Multiple interfaces, single code base
Clients ask CAS to run “actions” on data

Visual Interfaces

Programming Interfaces

API Interfaces

CAS Server
Open Architecture Advantage

• Same CAS action regardless of the interface and API.
• Results will be equivalent (given that all random elements are seeded).
• Information can be passed between different languages.
  • Promote tables to persist on the server and apply CAS actions from different APIs.
Interfaces
Building a Model from Scratch in the Visual Reporting Interface
Interfaces
Building a Model Using SAS Studio Snippets
Interfaces

Building a Model Using Open Source

1. Set up (same as previous exercise)

```python
import sas
from sas import *
from pprint import pprint
from matplotlib import pyplot as plt
import pandas as pd
import sys
%matplotlib inline

cashost='gatekrbhdps1.gatehadoop.com'
casport=5570
casauth='/home/sasdemo/.authinfo'
indata_dir='/opt/sas/data'
indata='hmeq'
sess = sas.CAS(cashost, casport, casauth=casauth, caslib='mycaslib')
sess.set_option(caslib='mycaslib')
results=sess.table.caslibInfo(caslib='mycaslib', verbose=True)
info=sess.table.fileinfo(caslib='mycaslib')
res=sess.table.loadtable(path='Hmeq.sas7bdat', caslib='mycaslib', casout='hmeq')
```
The New Languages of SAS 9.4 and SAS Viya: A SAS Programmer's Primer
SAS® Studio

SAS Studio is a browser-based programming and code-generation interface. It’s available via a browser on any device that connects to your SAS environment.
SAS® Studio
How to open

• Click on the Top Left action menu.

• Choose the first action, Develop SAS Code
SAS® Studio

Tips

• Tasks will generate any code necessary to complete the desired task.
• You can even write your own custom tasks.
• Learn and “borrow” from the generated code if you wish to expand your programming knowledge.
• (Folders for tasks will vary depending on what products are licensed)
Programming Interfaces
CAS
CAS Actions

- CAS actions are the tools used to interact with data on the CAS server.
- CAS actions are wrappers for parallel processing algorithms.
- CAS actions can load data, transform data, compute statistics, perform analytics, and create output.

Python Functions ≡ SAS 9.4 Procedures ≡ CAS Actions
The Python SWAT Package

• Gives unique Python functions to perform licensed CAS actions. The functionality mimics the look and feel of Python syntax, making it easy for Python users to take advantage of CAS.

```
sess.decisionTree.forestTrain(
    table = dict(),
    target = string,
    inputs = value_list,
    nominals = value_list,
    ntree = int32,
    casOut = dict()
)
```
Simple CAS Actions Example

```
loadtable/path="datasources/cars.csv"
   importOptions={fileType="csv"},
   caslib="CASTestTmp";
run;

histogram result=hist_result/nBins=1
   table={name="datasources.CARS"
      caslib="CASTestTmp"
      varList={name="sales"}};
run;
```
PROC versus CAS Action

```sas
proc factmac data=mycas.movlens nfactors=10 learnstep=0.15 maxiter=20 outmodel=mycas.factors;
   input userid itemid /level=nominal;
   target rating /level=interval;
   output out=mycas.out1 copyvars=(userid itemid rating);
run;
```

```sas
proc cas;
   action factmac result=R / table={name="movlens"},
      outModel={name="factors_out", replace=true},
      inputs={"userid", "itemid"},
      nominals={"userid", "itemid"},
      target="rating",
      maxIter=20, nFactors=10, learnStep=0.15,
      output={casout={name="score_out", replace="TRUE"},
               copyvars={"userid","itemid","rating"}};
run;
```
Open Source Interface

Jupyter Notebooks
Open Access between SAS, Python & R
SAS® Scripting Wrapper for Analytics Transfer (SWAT)

- Access to SAS® Viya™ from Python and R
- Integration of SAS® Analytics in Python and R code
- R Studio and Jupyter Notebook support
- Issue tracking and collaboration in development through GitHub project
proc print data = x.hmeq (obs = 10);
run;

df = s.CASTable('hmeq')
df.head(10)

df <- defCasTable(s, 'hmeq')
head(df, 10)

[table.fetch]
  table.name = "hmeq"
  from = 1 to = 10
Classification
Our example today

- The dataset is from a financial institution with customer demographics and loan/credit behavior.
- The goal of this modeling exercise is to predict which people are likely to default on a home equity loan.
- The data are at the customer-level (subject-level).
- n=5960
- columns = 13

<table>
<thead>
<tr>
<th>Alphabet List of Variables and Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>
# Home Equity

## Sample Data

<table>
<thead>
<tr>
<th>BAD</th>
<th>LOAN</th>
<th>MORTDUE</th>
<th>VALUE</th>
<th>REASON</th>
<th>JOB</th>
<th>YOJ</th>
<th>DEROG</th>
<th>DELINQ</th>
<th>CLAGE</th>
<th>NINQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1100</td>
<td>25860</td>
<td>39025</td>
<td>Homeimpr</td>
<td>Other</td>
<td>10.5</td>
<td>0</td>
<td>0</td>
<td>94.3666...</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1300</td>
<td>70053</td>
<td>68400</td>
<td>Homeimpr</td>
<td>Other</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>121.833...</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1500</td>
<td>13500</td>
<td>16700</td>
<td>Homeimpr</td>
<td>Other</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>149.466...</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1500</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>0</td>
<td>1700</td>
<td>97800</td>
<td>112...</td>
<td>Homeimpr</td>
<td>Office</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>93.333...</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1700</td>
<td>30548</td>
<td>40320</td>
<td>Homeimpr</td>
<td>Other</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>101.466...</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1800</td>
<td>48649</td>
<td>57037</td>
<td>Homeimpr</td>
<td>Other</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>77.1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1800</td>
<td>28502</td>
<td>43034</td>
<td>Homeimpr</td>
<td>Other</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>88.7660...</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2000</td>
<td>32700</td>
<td>46740</td>
<td>Homeimpr</td>
<td>Other</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>216.933...</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2000</td>
<td>.</td>
<td>62250</td>
<td>Homeimpr</td>
<td>Sales</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>115.8</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2000</td>
<td>22608</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>18</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>2000</td>
<td>20627</td>
<td>29800</td>
<td>Homeimpr</td>
<td>Office</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>122.533...</td>
<td>1</td>
</tr>
</tbody>
</table>
# Home Equity Profile

<table>
<thead>
<tr>
<th>Column</th>
<th>Unique</th>
<th>Null</th>
<th>Blank</th>
<th>Pattern Count</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAD</td>
<td>0% (2)</td>
<td>0</td>
<td></td>
<td></td>
<td>0.20</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
</tr>
<tr>
<td>CLAGE</td>
<td>94% (5314)</td>
<td>5% (308)</td>
<td></td>
<td></td>
<td>179.77</td>
<td>173.47</td>
<td>0.00</td>
<td>85.81</td>
</tr>
<tr>
<td>CLNO</td>
<td>1% (62)</td>
<td></td>
<td></td>
<td></td>
<td>21.30</td>
<td>20.00</td>
<td>16.00</td>
<td>10.14</td>
</tr>
<tr>
<td>DEBTINC</td>
<td>100% (4695)</td>
<td>21% (1267)</td>
<td></td>
<td></td>
<td>33.78</td>
<td>34.82</td>
<td>0.00</td>
<td>8.60</td>
</tr>
<tr>
<td>DELINQ</td>
<td>0% (14)</td>
<td></td>
<td></td>
<td></td>
<td>0.45</td>
<td>0.00</td>
<td>0.00</td>
<td>1.13</td>
</tr>
<tr>
<td>DEROG</td>
<td>0% (11)</td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.85</td>
</tr>
<tr>
<td>JOB</td>
<td>0% (6)</td>
<td></td>
<td></td>
<td></td>
<td>0% (279)</td>
<td>0</td>
<td>5</td>
<td>Other</td>
</tr>
<tr>
<td>LOAN</td>
<td>9% (540)</td>
<td>0</td>
<td></td>
<td></td>
<td>18,60...</td>
<td>16,30...</td>
<td>15,00...</td>
<td>11,20...</td>
</tr>
<tr>
<td>MORTDUE</td>
<td>92% (5053)</td>
<td>8% (515)</td>
<td></td>
<td></td>
<td>73.76...</td>
<td>65,01...</td>
<td>42,00...</td>
<td>44,45...</td>
</tr>
<tr>
<td>NINO</td>
<td>0% (16)</td>
<td></td>
<td></td>
<td></td>
<td>1.19</td>
<td>1.00</td>
<td>0.00</td>
<td>1.73</td>
</tr>
<tr>
<td>REASON</td>
<td>0% (2)</td>
<td></td>
<td></td>
<td></td>
<td>0% (252)</td>
<td>0</td>
<td>1</td>
<td>DebtC...</td>
</tr>
<tr>
<td>VALUE</td>
<td>92% (5381)</td>
<td>1% (112)</td>
<td></td>
<td></td>
<td>101,7...</td>
<td>89,23...</td>
<td>60,00...</td>
<td>57,38...</td>
</tr>
<tr>
<td>YOJ</td>
<td>1% (99)</td>
<td></td>
<td></td>
<td></td>
<td>8.92</td>
<td>7.00</td>
<td>0.00</td>
<td>7.57</td>
</tr>
</tbody>
</table>
SAS Viya Introduction

Demo
SAS Visual Analytics

Introduction
SAS® Visual Analytics

- Data Preparation
- Data Exploration & Analytics
- Interactive Reporting
- Governance
- Collaboration & Info Sharing
- Location Analytics
Data Preparation

- Access to diverse data
- Data quality functions
- Table and column profiling
- Filter & transform (append, join, transpose, etc.)
- View lineage
Visual Exploration

• Discover relationships, trends, outliers, clusters
• 3rd-party visualizations (e.g. D3, Google Chart)
• Forecasting and scenario analysis
• Decision trees
• Text analysis (e.g. word cloud)
Interactive Reporting

- Responsive and precise layouts
- Dashboard creation
- Report formatting for user interactivity; filters, prompts, linking, etc.
- Share, interact and collaborate
Location Analytics

- Travel-time analysis
- Travel-distance analysis
- Custom polygons or shapes
- Geographic enrichment
- Geographic clustering
Governance

- Identity Management (e.g. users, groups)
- Authorization (e.g. data, content, capabilities)
- Monitoring Performance
- Auditing
## Baseline VA PROCs and Action Sets

<table>
<thead>
<tr>
<th>Baseline VA</th>
<th>Action Set</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Step</td>
<td>dataStep</td>
<td>runCode</td>
</tr>
<tr>
<td>TRANSPOSE</td>
<td>transpose</td>
<td>transpose</td>
</tr>
<tr>
<td>DS2</td>
<td>ds2</td>
<td>runDS2</td>
</tr>
<tr>
<td>FEDSQL</td>
<td>fedSql</td>
<td>execDirect</td>
</tr>
<tr>
<td>TREESPLIT</td>
<td>decisionTree</td>
<td>dtreeCode; dtreeMerge; dtreePrune; dtreeScore; dtreeSplit; dtreeTrain; tuneDecisionTree; Srsact</td>
</tr>
<tr>
<td></td>
<td>optminer</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sampling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>table</td>
<td></td>
</tr>
<tr>
<td>CAS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SAS Visual Statistics

Introduction
SAS® Visual Statistics

Visual “drag & drop” Interface

Programming Interface

Data Preparation

Visual Exploration

Descriptive Analytics

Predictive Analytics

Assess and Deploy
Data Preparation

- Access to different data sources
- Training-Validation Data Partitioning
- Feature Engineering (e.g. parameters, interactions)
- Variable selection and missing values
Visual Exploration

- Interactively discover relationships, trends, outliers
- Smart autocharting
- Analytics driven visualizations
- Explore predicted outputs
Descriptive Analytics

- Summary statistics
- K-means Clustering & Parallel Coordinate Plots
- Correlation matrices, scatter plots, box plots, etc. to visualize relationships & findings
Predictive Analytics

- Linear Regression
- Logistic Regression & Non-parametric Logistic Regression
- GLM Regression
- Generalized Additive Model
- Decision Tree
- Group-By Processing
Access and Deploy

- Model comparison summaries
- Interactively assess models
- Assessment charts for partitioned data
- Model recipe to score new data
Introduction: Software Overview

Key Features of SAS® Visual Statistics 8.3

- Modeling Techniques
- (Visual Interface)
  - Clustering (k-means)
  - Linear Regression
  - Logistic Regression
  - GLM Regression
  - Decision Trees
- Common Features
  - Training-validation partitioning
  - Variable Importance / Profile
  - Model Assessment
  - Model comparison
  - Derivation of predictive outputs
  - Ability to export model statistics into Excel
  - Score Code
SAS® Visual Statistics
SAS Studio – Available Programming Tasks
SAS® Visual Statistics
Capabilities via Programming Interfaces

Common

- ASSESS (Assess Supervised Models)
- BINNING (Variable Binning)
- CARDINALITY (Cardinality Analysis)
- PARTITION (Sampling and Partitioning)
- VARIMPUTE (Missing Value Imputation)
- VARREDUCE (Variable Reduction)
- FREQTAB (Frequency & Cross Tabulation)
- CORRELATION (Correlation)
### Analytical Capabilities via Programming Interfaces

**GENSELECT** (Generalized Linear Model)  
**KCLUS** (Kmeans and Kmodes Clustering)  
**LOGSELECT** (Logistic Regression)  
**NLMOD** (Non-linear Regression)  
**PCA** (Principle Component Analysis)  
**REGSELECT** (Linear Regression)  
**PLSMOD** (Partial Least Square)  
**GAMMOD** (Generalized Additive Model)  

**Analytical**

- ICA (Independent Component Analysis)  
- MBC (Multivariate Gaussian)  
- MODELMATRIX (Design Matrix)  
- LMIxED (Linear Mixed Models)  
- TREESPLIT (Decision Trees)  
- QTRSELECT (Quantile Regression)  
- PHSELECT (Proportional Hazard Models)  
- SPC (Statistical Process Control)
### SAS® Visual Statistics

#### Action Sets

<table>
<thead>
<tr>
<th>Feature</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardinality Analysis</td>
<td>Partial Least Squares</td>
</tr>
<tr>
<td>Clustering</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>Dimension Reduction</td>
<td>Proportional Hazards Regression Modeling</td>
</tr>
<tr>
<td>Frequency and Cross Tabulation</td>
<td>Quantile Regression Modeling</td>
</tr>
<tr>
<td>Generalized Additive Models</td>
<td>Regression</td>
</tr>
<tr>
<td>Independent Component Analysis</td>
<td>Sampling and Partitioning</td>
</tr>
<tr>
<td>Mixed Modeling</td>
<td>Statistical Process Control</td>
</tr>
<tr>
<td>Model-Based Clustering</td>
<td></td>
</tr>
</tbody>
</table>
SAS Visual Statistics

Demo
SAS Visual Data Mining and Machine Learning

Introduction
SAS® Visual Data Mining and Machine Learning

Visual "drag & drop" Interface

Programming Interface

Data Preparation

Visual Exploration

Model Studio

Model Deployment

Machine Learning

Copyright © SAS Institute Inc. All rights reserved.
Data Preparation

- Access to different data sources
- Training-Validation Data Partitioning
- Feature Engineering (e.g. parameters, interactions)
- Variable selection and missing values
Visual Exploration

- Interactively discover relationships, trends, outliers
- Smart autocharting
- Analytics driven visualizations
- Explore predicted outputs
- Variable transformation
Model Studio

- Pipeline of activities
- Drag and drop and access to code
- Nodes are run asynchronously
- Reproducibility
- SAS best practice toolkit
Modern Machine Learning

- Forest
- Neural Network (including Deep Learning)
- Gradient Boosting
- Support Vector Machines
- Factorization Machines
- Bayesian Networks
- Autotuning
Comparison and Deploy

- Model comparison summaries
- Interactively assess models
- Assessment charts for partitioned data
- Publish score code; batch, API call, in-database
SAS® Visual Data Mining and Machine Learning 8.3
Visual Interface

Machine Learning Techniques
- Forest
- Factorization Machine
- Gradient Boosting
- Neural Network
- Support Vector Machine

Common Features
- Training-Validation
- Model Assessment
- Model Comparison
- Score Code or Astore Table
- Ability to export model statistics into Excel
SAS® Visual Data Mining and Machine Learning
SAS Studio – Available Programming Tasks
### SAS® Visual Data Mining and Machine Learning

#### Capabilities via Programming Interfaces

<table>
<thead>
<tr>
<th>Analytical</th>
<th>Analytical</th>
</tr>
</thead>
<tbody>
<tr>
<td>- FOREST (Random Forest Model)</td>
<td>- SVDD (Support Vector Data Description)</td>
</tr>
<tr>
<td>- GRADBOOST (Gradient Boosting)</td>
<td>- MWPCA (Moving Window Principal Component Analysis)</td>
</tr>
<tr>
<td>- NNET (Neural Network)</td>
<td>- RPCA (Robust Principal Component Analysis)</td>
</tr>
<tr>
<td>- SVMACHINE (Support Vector Machine)</td>
<td>- TSNE (t-distributed stochastic neighbor embedding)</td>
</tr>
<tr>
<td>- FACTMAC (Factorization Machine)</td>
<td>- ASTORE (Analytic Store for Models)</td>
</tr>
<tr>
<td>- NETWORK (Network analytics and community detection)</td>
<td>- FISM (Frequent Item Set)</td>
</tr>
<tr>
<td>- TEXTMINE (NLP and statistical analysis of text)</td>
<td>- MBANALYSIS (Market Basket Analysis)</td>
</tr>
<tr>
<td>- TMSCORE (Scores textural data)</td>
<td>- GVARCLUS (Graphical Variable Clustering)</td>
</tr>
<tr>
<td>- BOOLRULE (Boolean rules extraction)</td>
<td>- FASTKNN (K- Nearest Neighbor)</td>
</tr>
<tr>
<td>- MTLEARN (Multi Learning task for least squares loss)</td>
<td>- BNET (Bayesian Network)</td>
</tr>
<tr>
<td>- SEMISUPLEARN (graph-based semisupervised learning algorithm)</td>
<td>- Deep Learning (no procedures, through CAS actions)</td>
</tr>
</tbody>
</table>

Copyright © SAS Institute Inc. All rights reserved.
<table>
<thead>
<tr>
<th>Action Sets</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytic Store Scoring</td>
<td>Image</td>
</tr>
<tr>
<td>Association Rule Mining</td>
<td>Language Model</td>
</tr>
<tr>
<td>Audio Action</td>
<td>MLTools</td>
</tr>
<tr>
<td>Autotune</td>
<td>Network</td>
</tr>
<tr>
<td>Bayesian Net Classifier</td>
<td>Neural Network</td>
</tr>
<tr>
<td>BioMedImage</td>
<td>Nonparametric Bayes</td>
</tr>
<tr>
<td>Boolean Rule</td>
<td>Robust PCA</td>
</tr>
<tr>
<td>Factorization Machine</td>
<td>Support Vector Data Description</td>
</tr>
<tr>
<td>Fast k-Nearest Neighbors Algorithm</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>Generalized Linear Multitask Learning</td>
<td>Text Mining</td>
</tr>
<tr>
<td>Graph-Based Semisupervised Learning</td>
<td>TSNE</td>
</tr>
<tr>
<td>Graphical Variable Clustering</td>
<td></td>
</tr>
</tbody>
</table>
SAS Visual Data Mining and Machine Learning

Visual Interface Demo
SAS Visual Text Analytics
Unearthing the full potential within complex data sources can be tricky

<table>
<thead>
<tr>
<th>Language is messy!</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Large data volumes and inconsistent formats</td>
</tr>
<tr>
<td>• Multiple sources and languages</td>
</tr>
<tr>
<td>• Misspellings, slang, and abbreviations</td>
</tr>
<tr>
<td>• Highly subjective to interpretation and context</td>
</tr>
</tbody>
</table>

Manual review is both inconsistent and time consuming, and a sampling approach can mean missing out on valuable information and the big picture.
SAS is continually extending its Natural Language Processing capabilities

Natural Language Processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret and manipulate human language.

- Parsing and Entity/Relationship Extraction
- Topic Detection, Text Clustering & Profiling
- Natural Language Understanding & Generation
- Classification (Categories, Sentiment)
- Automatic Summarization, Search
- Auto-Translation* & Speech to Text

* An area of NLP not currently offered by SAS
**SAS Natural language processing at a glance**

SAS Text Analytics solutions perform **document conversion**, **tokenization**, **lemmatization (stemming)**, **part-of-speech detection**, and **apply lexicons** for misspellings, synonyms, multi-word terms, and start/stop words.

<table>
<thead>
<tr>
<th>Import Data</th>
<th>I disputed a crediter for the incorrect amt they reported to Trans Union credit bureau.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parse and Detect Entities</td>
<td>I disputed a crediter for the incorrect amt they reported to Trans Union credit bureau.</td>
</tr>
<tr>
<td>Filter, Tag, and Resolve Terms</td>
<td>I disputed a crediter for the incorrect amt they reported to Trans Union credit bureau.</td>
</tr>
</tbody>
</table>

- Automatically detect term variants, misspellings and parts-of-speech
- Resolve shorthand, slang, and abbreviations with synonym lists
<table>
<thead>
<tr>
<th>Language</th>
<th>Language</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>Greek</td>
<td>Romanian</td>
</tr>
<tr>
<td>Chinese</td>
<td>Hebrew</td>
<td>Russian</td>
</tr>
<tr>
<td>Croatian</td>
<td>Hungarian</td>
<td>Slovak</td>
</tr>
<tr>
<td>Czech</td>
<td>Hindi</td>
<td>Slovene</td>
</tr>
<tr>
<td>Danish</td>
<td>Indonesian</td>
<td>Spanish</td>
</tr>
<tr>
<td>Dutch</td>
<td>Italian</td>
<td>Swedish</td>
</tr>
<tr>
<td>English</td>
<td>Japanese</td>
<td>Tagalog</td>
</tr>
<tr>
<td>Farsi</td>
<td>Korean</td>
<td>Thai</td>
</tr>
<tr>
<td>Finnish</td>
<td>Norwegian</td>
<td>Turkish</td>
</tr>
<tr>
<td>French</td>
<td>Polish</td>
<td>Vietnamese</td>
</tr>
<tr>
<td>German</td>
<td>Portuguese</td>
<td></td>
</tr>
</tbody>
</table>
SAS® Visual Text Analytics
Visual Interface

Text Analytics Techniques
• Contextual Extraction
• Text Parsing
• Sentiment Analysis
• Develop Topics
• Content Categorization

Common Features
• Report ready output
• Model ready output
• Scoring
Visual Pipeline

- Customizable and portable nodes and pipelines
- Quick save of output data from Topics and Categories
- Easy access to automatically-generated score code
- Native integration and common interface with data preparation, visualization, model management, data lineage, and other analytics solutions
Inspect and Manage Terms

• View Kept/Dropped terms and their variants, roles, and frequency in your document collection

• Explore Term Maps and Similar Terms

• Review matched documents to see terms used in context
Explore and Manage Discovered Topics

- Text Topics are discovered automatically, without requiring explicit input from the user.
- View the terms and document comprising a topic and how pervasive each topic is in the collection.
- Merge or split topics, create custom topics from selected terms, or promote topics to become rule-based categories.
Content Categorization

- Build, test, and assess category definitions and manage your taxonomy structure
- Leverage machine-built rules, or author your own definitions using keywords and a broad range of Boolean and linguistic operators & qualifiers
- Use Textual Elements and the rule-builder to help write rules from scratch
- Produce report-ready or modeling-ready categorized output
Contextual Extraction

• Leverage pre-defined entities for standard concepts such as dates, locations, and measures

• Create custom entities using a broad range of rule types, operators, regular expressions, and linguistic qualifiers

• Extract custom facts (relationships between entities)
Sentiment Analysis

• Create custom entities using a broad range of rule types, operators, regular expressions, and linguistic qualifiers

• Extract custom facts (relationships between entities)

Available languages for domain-independent sentiment model

- Arabic
- Chinese (Simplified)
- Chinese (Traditional)
- Dutch
- English
- Farsi
- French
- German
- Italian
- Japanese
- Korean
- Portuguese
- Spanish
- Turkish
SAS Studio
VTA Pre-built tasks

- Wizard-driven tasks
- Options to use algorithms not presented in Model Studio GUI
- Easily access intermediate and output tables
- Auto-generate SAS code for further development
SAS® Visual Text Analytics
Capabilities via Programming Interfaces

Analytical

- TEXTMINE (NLP and statistical analysis of text)
- TMSCORE (Scores textural data)
- BOOLRULE (Boolean rules extraction)
<table>
<thead>
<tr>
<th>Action Set</th>
<th>Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boolean Rule</td>
<td>Text Mining</td>
</tr>
<tr>
<td>Deep RNN</td>
<td>Text Parse</td>
</tr>
<tr>
<td>LDA Topic Modeling</td>
<td>Text Rule Development</td>
</tr>
<tr>
<td>Search</td>
<td>Text Rule Discovery</td>
</tr>
<tr>
<td>Search Analytics</td>
<td>Text Rule Score</td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>Text Summarization</td>
</tr>
<tr>
<td>Smart Data Set Analysis</td>
<td>Text Topics</td>
</tr>
<tr>
<td>TA Conditional Random Fields</td>
<td>Text Utilities</td>
</tr>
</tbody>
</table>
SAS Visual Text Analytics
Open Source

Score Categories

In [14]:
   s.textRuleScore.applyCategory(
   model={"caslib": mco_binary_caslib, "name": mco_binary_table_name},
   table={"caslib": inputTableLib, "name": inputTableName},
   docId=keyColumn,
   text=textColumn,
   casOut={"caslib": clName, "name": outCategories, "replace": True},
   matchOut={"caslib": clName, "name": outMatches, "replace": True}
)

Out[14]:

$OutputCategoricalFull$

<table>
<thead>
<tr>
<th>casLib</th>
<th>Name</th>
<th>Label</th>
<th>Rows</th>
<th>Columns</th>
<th>casTable</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASUSER(thsia)</td>
<td>Out_Categories</td>
<td>65564</td>
<td>4</td>
<td>casTable(Out_Categories)</td>
<td>caslib=CASUSER(thsia)</td>
</tr>
<tr>
<td>CASUSER(thsia)</td>
<td>Out_TermMatches</td>
<td>109034</td>
<td>5</td>
<td>casTable(Out_TermMatches)</td>
<td>caslib=CASUSER(thsia)</td>
</tr>
</tbody>
</table>

elapsed 2.92s · user 9.88s · sys 2.99s · mem 274MB

Review Outputs

In [16]:
   s.CASTable(outCategories).head(10)

Out[16]:

Selected Rows from Table OUT_CATEGORIES

<table>
<thead>
<tr>
<th>Unique_ID</th>
<th>_result_id</th>
<th><em>category</em></th>
<th><em>score</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>157.0</td>
<td>HELPERS/Mastember</td>
<td>1.0</td>
</tr>
<tr>
<td>1</td>
<td>567.0</td>
<td>HELPERS/PosKeywordsPPL</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>457.0</td>
<td>HELPERS/PosKeywordsPPL</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>557.0</td>
<td>Atmosphere/Appearance/Clean_Atractive</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Consumer Complaints

Our example today

• The dataset is from the Consumer Financial Protection Bureau (CFPB) and contains complaints submitted by customers
• The goal of this exercise is to analyze the text data from their complaints
• The data are at the customer-level (subject-level)
• n=2.4K
• columns = 18

Narrative

1. My husband and I bought a car in XXXX North Carolina. We drove home to GA to find out that the car had so many problems, we broke down at the XXXX Car Haus to fix the car. After long combination ($500.00) of fixes, my car was still not fixed. I told my bank to write off the account balance on my account and I have to take it.
2. On XXXX XXXX, I transferred XXXX from my bank account to my checking account in an attempt to fix if was not working as it should. Before I had already made a verification with the XXXX card. The account was overdrawn and I was charged a late fee of $50.00. My bank also sent me a letter saying I was late paying the account. My bank account was overdrawn by $50.00 and I was charged a late fee. I had a search of my account for my own and there was no record.
3. A check was deposited in my account by the bank without my consent. The bank did not send me a confirmation letter for my own. My account was closed over this by Early Warning Services LLC. This company will not take me off the list.

Products

- Closed with explanation
- Closed with non-monetary relief
- Closed with monetary relief
- Unresolved
Consumer Complaints Details
# Consumer Complaints

## Sample Data

![Image of SAS software interface showing Consumer Complaints sample data]

<table>
<thead>
<tr>
<th>date_received</th>
<th>Product</th>
<th>Sub_prod</th>
<th>Issue</th>
<th>Sub_issue</th>
<th>Consumer_complaint_narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>06/19/2...</td>
<td>Bank ac...</td>
<td>Checkin...</td>
<td>Deposit...</td>
<td></td>
<td>I live paycheck to paycheck. I check my bank account balance daily if not XXXX a ... Compl</td>
</tr>
<tr>
<td>06/19/2...</td>
<td>Bank ac...</td>
<td>Checkin...</td>
<td>Deposit...</td>
<td></td>
<td>I have been a customer with Wells Fargo for nearly 20 years – we have our checkin... Compl</td>
</tr>
<tr>
<td>06/17/2...</td>
<td>Bank ac...</td>
<td>Other b...</td>
<td>Making/...</td>
<td></td>
<td>To Whom It May Concern: Dear Sir/Madam needed to pay XXXX AED to a lawyer... Compl</td>
</tr>
<tr>
<td>06/17/2...</td>
<td>Credit c...</td>
<td>Other</td>
<td>Other</td>
<td></td>
<td>Without my permission, Wells Fargo took money from my bank account to pay for ... Compl</td>
</tr>
<tr>
<td>06/16/2...</td>
<td>Credit c...</td>
<td>Other</td>
<td>Other</td>
<td></td>
<td>I opened a secured credit card through Wells Fargo in XXXX of 2013 in the amount... Compl</td>
</tr>
<tr>
<td>06/16/2...</td>
<td>Mortgage</td>
<td>FHA mo...</td>
<td>Loan m...</td>
<td></td>
<td>I purchased a Condo in XXXX in XXXX the mortgage was for ($1400.00) monthly ... Compl</td>
</tr>
<tr>
<td>06/15/2...</td>
<td>Bank ac...</td>
<td>Checkin...</td>
<td>Making/...</td>
<td></td>
<td>On XXXX/XXXX/2015 my account was negative ($170.00) and some change. I do... Compl</td>
</tr>
<tr>
<td>06/14/2...</td>
<td>Debt co...</td>
<td>Non-fed...</td>
<td>Cont’d...</td>
<td></td>
<td>I paid of my JP Morgan Chase student loan more than XXXX years ago and even r... Compl</td>
</tr>
<tr>
<td>06/13/2...</td>
<td>Bank ac...</td>
<td>Other b...</td>
<td>Making/...</td>
<td></td>
<td>1. I authorize withdrawals from my checking account for my payment and they do ... Compl</td>
</tr>
<tr>
<td>06/13/2...</td>
<td>Credit c...</td>
<td>Late fee</td>
<td>Late fee</td>
<td></td>
<td>My credit card company charged me ($26.00) for late fee on a ($20.00) interest at ... Compl</td>
</tr>
<tr>
<td>06/13/2...</td>
<td>Debt co...</td>
<td>Credit c...</td>
<td>Commu...</td>
<td></td>
<td>Citi Bank has called me a minimum of 3 times (sometimes up to XXXX) every day ... Compl</td>
</tr>
<tr>
<td>06/12/2...</td>
<td>Credit c...</td>
<td>Credit d...</td>
<td>Credit d...</td>
<td></td>
<td>I applied online for a credit card with Chase Bank USA, N.A. XXXX XXXX XXXX, X... Compl</td>
</tr>
<tr>
<td>06/12/2...</td>
<td>Mortgage</td>
<td>Convent...</td>
<td>Loan m...</td>
<td></td>
<td>We have had an extremely difficult time with our mortgage company, CitiMortgage... Compl</td>
</tr>
<tr>
<td>06/12/2...</td>
<td>Mortgage</td>
<td>Convent...</td>
<td>Loan m...</td>
<td></td>
<td>I started a loan modification with Citi mortgage in 2011 which was supposed to co... Compl</td>
</tr>
<tr>
<td>06/11/2...</td>
<td>Bank ac...</td>
<td>Checkin...</td>
<td>Deposit...</td>
<td></td>
<td>Our sons and daughter send us money XXXX XXXX (Thursday) and XXXX XXXX ... Compl</td>
</tr>
<tr>
<td>06/11/2...</td>
<td>Bank ac...</td>
<td>Other b...</td>
<td>Account...</td>
<td></td>
<td>I feel that my bank chase is withdrawing money or fees for insufficient funds ... Compl</td>
</tr>
<tr>
<td>06/10/2...</td>
<td>Mortgag...</td>
<td>Franchise</td>
<td>Winnow</td>
<td></td>
<td>Wells Fargo on direct deposit advences they stopped such transactions and ran ... Compl</td>
</tr>
</tbody>
</table>
Consumer Complaints
Profile
Narratives

Consumer Complaints Submitted to CFPB
Categorization
Documents are categorized as related to Products, Overdrafts, Legal Issues or Digital Currency.
Sentiment

Do people feel positive, negative or neutral toward something?
### Concept Extraction
Extract specifics, like interest rate, from the text.

<table>
<thead>
<tr>
<th>Concept</th>
<th>Matched Text</th>
<th>Narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest Rate</td>
<td>4.25%</td>
<td>On XXXX XXXX I was contacted by Royal United Mortgage via a request I submitted via XXXX XXXX. At the time XXXX X... We are filing a complaint regarding a situation that resulted in our paying an additional [2100.00] in exchange fees... I signed up for an american express everyday credit card under the premise that my balance transfer would be 0% interest.</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>7%</td>
<td>On XXXX XXXX I received an incentive email from Bank of America regarding opening up a credit card. Upon receiving this offer, I responded to an offer from Bank of America for its &quot;Travel Rewards&quot; card. The offer included 0% interest for 12 months...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>0%</td>
<td>On XXXX XXXX I applied for a credit card with XXXX XXXX XXXX online with the offer a balance transfer with 0% interest...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>0%</td>
<td>In XXXX 2017 I applied online for the XXXX XXXX XXXX XXXX and was approved along with a 7 month 0%/APR I was...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>1%</td>
<td>In 2017, Best buy charged me the full interest for 2 item amounts that were interest deferred. Total: XXXX $. All interest was...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>1.5%</td>
<td>On XXXX XXXX, I attempted to re-finance my XXXX XXXX XXXX at XXXX XXXX XXXX in XXXX, CA. I was doing this so...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>1.5%</td>
<td>On XXXX XXXX XXXX, I attempted to re-finance my XXXX XXXX XXXX at XXXX XXXX XXXX in XXXX, CA. I was doing this...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>1%</td>
<td>On XXXX XXXX XXXX, I attempted to re-finance my XXXX XXXX XXXX at XXXX XXXX XXXX in XXXX, CA. I was doing this...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>10%</td>
<td>I assumed the mortgage from XXXX XXXX XXXX I have made the payments of [5570.00] after the assumption of...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>13.99%</td>
<td>On XXXX XXXX, I applied for a credit card with XXXX XXXX XXXX online with the offer a balance transfer with 0% interest...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>17%</td>
<td>As many filed complaints against XXXX XXXX XXXX, why this vehicle lender doesn't want to reduce my interest of 17% o...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>19.99%</td>
<td>I obtained a line of credit from Paypal in the beginning of XXXX to make a singular, particular purchase. Paypal offers a pr...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>2%</td>
<td>I have made several requests to get this HomeEquityLine ofCreditRefinanced. In XX/XX/XX/XX I filed a complaint with...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>2.8%</td>
<td>I took out a credit card from Capital One several years ago after experiencing a XXXX illness which left me XXXX XXXX XXXX...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>24%</td>
<td>I took out a credit card from Capital One several years ago after experiencing a XXXX illness which left me XXXX XXXX XXXX...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>25.9%</td>
<td>Attn: GD/CareCredit/Synchrony Bank/Consumer Financial Protection Bureau. To Whom It May Concern: In regards to the...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>26.99%</td>
<td>XXXX XXXX XXXX (CFPB Case XXXX) Act ending in XXXX 117 nor XXXX XXXX was ever under any circumstances appropri...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>26.99%</td>
<td>XXXX XXXX XXXX XXXX (CFPB Case XXXX) Act ending in XXXX by nor XXXX XXXX was ever under any circumstance...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>27%</td>
<td>I went into XXXX XXXX in XX/XX/XX in or XXXX XXXX to get a personal loan, and was approved. Over the remain...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>29.9%</td>
<td>I am writing to report unexplained fees charged on my XXXX XXXX XXXX XXXX credit card. It is my financial habit to...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>29.9%</td>
<td>In XXXX/XXX, I was in the midst of a financial crisis, losing my job and having to relocate and downsizes. At that time I...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>3.5%</td>
<td>I contacted Milend Inc in XXXX, GA back in XXXX of 2017 of securing a mortgage through them. Spoke with a young girl...</td>
</tr>
<tr>
<td>Interest Rate</td>
<td>3.5%</td>
<td>I had a private loan through XXXX XXXX XXXX in which averaged at 2.8% interest rate. A few months ago I was notified...</td>
</tr>
</tbody>
</table>
### Topic Discovery

**Which topics appear in my document collection?**

<table>
<thead>
<tr>
<th>Term</th>
<th>Score</th>
<th>Narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td>loan</td>
<td>1.0</td>
<td>Caliber home loans - I was mislead by the amount</td>
</tr>
<tr>
<td>+mortgage.</td>
<td>1.0</td>
<td>After graduating I looked for options</td>
</tr>
<tr>
<td>+modification.</td>
<td>1.0</td>
<td>to consolidate my student loans. I was able</td>
</tr>
<tr>
<td>+payment.</td>
<td>1.0</td>
<td>to do this with all outstanding loans except</td>
</tr>
<tr>
<td>+home.</td>
<td>1.0</td>
<td>Navient. I was told it cost</td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
<td>I can pay these loans either with my income</td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
<td>XXXXXX XXXXXX (XXXX university) &amp; XXXXXX XXXXXX (XXXX University) Loans have been transferred several</td>
</tr>
<tr>
<td>-0.1</td>
<td></td>
<td>times to different service...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>Navient has my loans. My loan payment amount continues to go up</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>by hundreds of dollars and this is really affecting my credit. Can you please help me. I tried...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX XXXXXX (FL, XXX University) would like to start this letter by stating that I am XXXXX years old and what h...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>Starting in the summer of 2015 I had wanted to consolidate all my student loans so I could have one single manageable payment. When I found a consolidation...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>My concern is that I was recently informed by Navient that I do not qualify for the Public Service Student Loan Forgiveness Program. I have been a XXXXXX XXXXXX...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>H, One more than one occasion, Navient has offered to lower my interest rate (my rates for my private loans are : 5.9, 6.1, 9.8 %). They do this over the p...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>In 2005, I was forced into taking a private, co-signed loan for around $6000.00 with XXXXXX (now navient). I was told to defer my loan or use for emergency...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>NAVIENT This loan was paid off. Reinvestigate and erase please.</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I have been in contact with Navient to attempt to get on IBR to repay my student loans, but I've been given the run around. I was on a $ XXXX/month payment...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>My house was sold in a foreclosure sale, while I was trying to work out a Workable Solution with Citimortgage. I contacted with the foreclosure team a Citimortg...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I have submitted financial requests for mortgage modification to Ocen Servicing several times. Ocwen is in willing to work with me to keep my home...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I have had my loan go into default a few years back and I have corrected it, but I am getting harassing never ending calls and emails about my loan will be gone...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>We have a problem with Fay Servicing who recently took over our families mortgage loans. As soon as they got this loan they set an auction date before informing...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>My student loans have increased from XXXX to XXXX XXXXX dollars and the school I attended was no where near that cost and I was in the military but so...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I had a private loan through XXXX XXXXX XXXXX in which averaged a 2.8 % interest rate. A few months ago I was notified that Navient will be taking over the loan...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>After speaking with Navient on XXXXXXXXX after several times, I am disappointed in their lack of helpfulness in allowing me to make student loan payments...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>For years I have reaching out about how to make payments to my Navient loan (s). I learned that payments were only made to select loans and I found myself...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>This is the second complaint I am filling for this company. The original complaint number is XXXXX. Originally I was working with Option Funding on refinancing...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>XXXXXXX XXXXX. Back in XXXXXXXX of XXXX, I was in the middle of an ugly divorce. My Ex Wife was responsible for all the monthly bills but she never paid a...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I RECEIVED A PACKAGE FROM CAPITAL ONE ON XXXXXXX XXXXX FOR BORROWER ASSISTANCE LOAN. I HAVE A HELLOC Loan THAT WOULD MATUR...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I moved in to my rental due to financial problem found out later it was identify theft. I started a home modification in XXXXXXX that took 5 months for the go-ahead. T...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>I have a student loan with Navient. Been paying them for a little over a year and they keep increasing my loan. That its affecting my credit score as well...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>PURCHASE VEHICLE CAR SALES MAN DIDN'T TELL ME HE WAS RUNNING CREDIT TO DIFFERENT AUTO FINANCE COMPANY...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>To Whom It May Concern: I am writing this letter to inform you of the deprorable customer service and the unfair and deceptive acts and/or practices of XXXXX...</td>
</tr>
<tr>
<td>+0.0</td>
<td></td>
<td>On several occasions I contacted Navient regarding my payment being too high. I am a XXXXXXX XXXXX and they told me O do not qualify for loan forgiven...</td>
</tr>
</tbody>
</table>
SAS Visual Text Analytics
Demo
## End-to-End Analytical Capabilities

### Data Manipulation
- In-Memory Data Step
- Frequency / Crosstab
- Data Transpose
- Variable Binning
- Variable Cardinality Analysis
- Variable Summary
- Sampling and Partitioning
- Missing Value Imputation
- Variable Selection
- Model Assessment
- DS2
- FedSQL
- S3 Data Connector

### Statistics
- Cox Proportional Hazards
- Decision Trees
- Design Matrix
- General Additive Models
- Generalized Linear Models
- Independent Component Analysis
- K-means and K-modes Clustering
- Linear Regression
- Linear Mixed Models
- Logistic Regression
- Model-Based Clustering
- Nonlinear Regression
- Ordinary Least Squares Regression
- Partial Least Squares Regression
- Pearson Correlation
- Principal Component Analysis
- Quantile Regression
- Shewhart Control Chart Analysis

### Machine Learning
- Bayesian Networks
- Boolean Rules
- Factorization Machines
- Frequent Item Set Mining
- Gaussian Mixture Model
- Gradient Boosting
- K Nearest Neighbor
- Image Processing (including Biomedical)
- Market Basket Analysis
- Moving Windows PCA
- Multitask Learning
- Network Analytics / Community Detection
- Neural Networks / Deep Learning
- Random Forest
- Robust PCA
- Semi-supervised Learning
- Support Vector Data Description
- Support Vector Machines
- t-distributed Stochastic Neighbor Embedding
- Text Mining
- Variable Clustering

### Text Analytics
- Word Cloud
- Topics
- Text Parsing
- Content Categorization
- Sentiment Analysis
- Contextual Extraction

*For documentation please click [here](#) *

*Optimization and Forecasting are also supported*
Resources
SAS® Viya Resources

Videos

• 6 minutes getting started video
• 5 minutes getting started video
• 8 minute demo video
• Feature Engineering video

Using the Automated Analysis Feature in SAS® Visual Analytics in SAS® Viya®

Getting Started with Data Mining and Machine Learning Pipelines on SAS® Viya®

Building and Using Pipelines in SAS® Visual Forecasting
SAS® Viya

Resources

http://support.sas.com/software/products/visual-statistics/index.html#s1=2

SAS Visual Data Mining and Machine Learning User’s Guide
http://support.sas.com/software/products/visual-data-mining-machine-learning/index.html#s1=1

SAS Visual Text Analytics User’s Guide
http://support.sas.com/software/products/visual-text-analytics/index.html

Overview, Training, Samples and Tips

• SAS Viya Overview
• SAS Viya Training
• A Beginner’s Guide to Programming in the SAS® Cloud Analytics Services Environment
Resources
Programming

- SAS Studio
- CAS actions documentation
- SAS Github page for SWAT-Python
- SAS Github page for SWAT-R
- More example scripts for using SWAT-R & SWAT-Python
Useful Websites
Developer.sas.com, Communities.sas.com
Questions?

Thank you for your time and attention!

Connect with me:
LinkedIn: https://www.linkedin.com/in/melodierush
Twitter: @Melodie_Rush