Table 1. Factor and Numeric Variables	
Variable	**Description**
gender	What is your gender?
age	What is your age (# years)?
country	In which country do you currently reside?
education	What is the highest level of formal education that you have attained or plan to attain within the next 2 years?
undergraduate_major	Which best describes your undergraduate major
job_role	Select the title most similar to your current role (or most recent title if retired)
industry	In what industry is your current employer/contract (or your most recent employer if retired)?
years_experience	How many years of experience do you have in your current role?
ML_atwork	Does your current employer incorporate machine learning methods into their business?
percent_actively.coding	Approximately what percent of your time at work or school is spent actively coding?
How.long.have.you.been.writing.code.to.analyze.data.	How long have you been writing code to analyze data?
For.how.many.years.have.you.used.machine.learning.methods..at.work.or.in.school..	For how many years have you used machine learning methods (at work or in school)?
Do.you.consider.yourself.to.be.a.data.scientist.	Do you consider yourself to be a data scientist?
Programming_language_used_most_often	What specific programming language do you use most often?
wage	What is your current yearly compensation (approximate $USD)?

NOTE: Table 2 provides an overview of the one-hot encoded variables included in the dataset. Namely, in the original survey, respondents were asked questions for which they could select more than one option. For example, on the question: **Select any activities that make up an important part of your role at work** , the respondents could have selected:

A - Analyze.and.understand.data.to.influence.product.or.business.decisions
B - Build.and.or.run.a.machine.learning.service.that.operationally.improves.my.product.or.workflows
C - Build.and.or.run.the.data.infrastructure.that.my.business.uses.for.storing..analyzing..and.operationalizing.data
D - Build.prototypes.to.explore.applying.machine.learning.to.new.areas
E - Do.research.that.advances.the.state.of.the.art.of.machine.learning
F - None.of.these.activities.are.an.important.part.of.my.role.at.work

In the dataset, each option (A, B, C. ...) is a seperate variable. Everytime the respondent selected an option, this will be marked with a 1 in the appropriate colum, and 0 otherwise. The same structure is maintained for all other questions in this group.

Table 2. One-hot encoded variables	
Variable	**Description**
Activities_Analyze.and.understand.data.to.influence.product.or.business.decisions	Select any activities that make up an important part of your role at work
Activities_Build.and.or.run.a.machine.learning.service.that.operationally.improves.my.product.or.workflows	Select any activities that make up an important part of your role at work
Activities_Build.and.or.run.the.data.infrastructure.that.my.business.uses.for.storing..analyzing..and.operationalizing.data	Select any activities that make up an important part of your role at work
Activities_Build.prototypes.to.explore.applying.machine.learning.to.new.areas	Select any activities that make up an important part of your role at work
Activities_Do.research.that.advances.the.state.of.the.art.of.machine.learning	Select any activities that make up an important part of your role at work
Activities_None.of.these.activities.are.an.important.part.of.my.role.at.work	Select any activities that make up an important part of your role at work
Notebooks_Kaggle.Kernels	Which of the following hosted notebooks have you used at work or school in the last 5 years?
Notebooks_Google.Colab	Which of the following hosted notebooks have you used at work or school in the last 5 years?
Notebooks_Azure.Notebook	Which of the following hosted notebooks have you used at work or school in the last 5 years?
Notebooks_Google.Cloud.Datalab	Which of the following hosted notebooks have you used at work or school in the last 5 years?
Notebooks_JupyterHub.Binder	Which of the following hosted notebooks have you used at work or school in the last 5 years?

Notebooks_None	Which of the following hosted notebooks have you used at work or school in the last 5 years?
cloud_Google.Cloud.Platform..GCP.	Which of the following cloud computing services have you used at work or school in the last 5 years?
cloud_Amazon.Web.Services..AWS.	Which of the following cloud computing services have you used at work or school in the last 5 years?
cloud_Microsoft.Azure	Which of the following cloud computing services have you used at work or school in the last 5 years?
cloud_IBM.Cloud	Which of the following cloud computing services have you used at work or school in the last 5 years?
cloud_Alibaba.Cloud	Which of the following cloud computing services have you used at work or school in the last 5 years?
cloud_I.have.not.used.any.cloud.providers	Which of the following cloud computing services have you used at work or school in the last 5 years?
Programming_Python	What programming languages do you use on a regular basis?
Programming_R	What programming languages do you use on a regular basis?
Programming_SQL	What programming languages do you use on a regular basis?
Programming_Bash	What programming languages do you use on a regular basis?
Programming_Java	What programming languages do you use on a regular basis?
Programming_Javascript.Typescript	What programming languages do you use on a regular basis?
Programming_Visual.Basic.VBA	What programming languages do you use on a regular basis?
Programming_C.C..	What programming languages do you use on a regular basis?
Programming_MATLAB	What programming languages do you use on a regular basis?
Programming_Scala	What programming languages do you use on a regular basis?
Programming_Julia	What programming languages do you use on a regular basis?
Programming_SAS.STATA	What programming languages do you use on a regular basis?
Programming_None	What programming languages do you use on a regular basis?
ML_framework_Scikit.Learn	What machine learning frameworks have you used in the past 5 years?
ML_framework_TensorFlow	What machine learning frameworks have you used in the past 5 years?
ML_framework_Keras	What machine learning frameworks have you used in the past 5 years?
ML_framework_PyTorch	What machine learning frameworks have you used in the past 5 years?
ML_framework_Spark.MLlib	What machine learning frameworks have you used in the past 5 years?
ML_framework_H20	What machine learning frameworks have you used in the past 5 years?
ML_framework_Caret	What machine learning frameworks have you used in the past 5 years?
ML_framework_Xgboost	What machine learning frameworks have you used in the past 5 years?
ML_framework_randomForest	What machine learning frameworks have you used in the past 5 years?
ML_framework_None	What machine learning frameworks have you used in the past 5 years?
Visualization_ggplot2	What data visualization libraries or tools have you used in the past 5 years?
Visualization_Matplotlib	What data visualization libraries or tools have you used in the past 5 years?
Visualization_Altair	What data visualization libraries or tools have you used in the past 5 years?
Visualization_Shiny	What data visualization libraries or tools have you used in the past 5 years?
Visualization_Plotly	What data visualization libraries or tools have you used in the past 5 years?
Visualization_None	What data visualization libraries or tools have you used in the past 5 years?
data_Categorical.Data	Which types of data do you currently interact with most often at work or school?
data_Genetic.Data	Which types of data do you currently interact with most often at work or school?

data_Geospatial.Data	Which types of data do you currently interact with most often at work or school?
data_Image.Data	Which types of data do you currently interact with most often at work or school?
data_Numerical.Data	Which types of data do you currently interact with most often at work or school?
data_Sensor.Data	Which types of data do you currently interact with most often at work or school?
data_Tabular.Data	Which types of data do you currently interact with most often at work or school?
data_text.Data	Which types of data do you currently interact with most often at work or school?
data_Time.Series.Data	Which types of data do you currently interact with most often at work or school?
data_Video.Data	Which types of data do you currently interact with most often at work or school?
explainability.model_Examine.individual.model.coefficients	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_examine.feature.correlations	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_Examine.feature.importances	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_Create.partial.dependence.plots	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_LIME.functions	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_SHAP.functions	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?
explainability.model_None.I.do.not.use.these.model.explanation.techniques	What methods do you prefer for explaining and/or interpreting decisions that are made by ML models?