# Alteryx Predictive Master

## Certification Exam Prep Guide

# Earning your Predictive Master Certification takes you to further...

The world is full of dirty data and complex problems. Luckily, there are analytic adventure seekers like you who want to explore all that Alteryx Designer holds. Whether you're brand-new or a seasoned veteran, the Alteryx Certification Program provides you with ways to continue learning and a place to prove your Alteryx analytic skills.


## The Predictive Master Certification exam is designed to help you prove your skills and set you apart.

## Contents

# Exam Overview

The Predictive Master Certification allows you to demonstrate your mastery of using predictive analytics and machine learning to provide data-driven insights to your organization. Predictive Masters have full command of the data science lifecycle, including data preparation, data exploration, predictive modeling and evaluation, and interpretation and deployment. The scope of this exam includes predictive analytics concepts and all tools available in the Predictive Tools installation.


**AUDIENCE:** Anyone who is currently Advanced Certified

**EXPERIENCE LEVEL:** Expert

**PRICE:** $150 USD

**EXAM TYPE:** User and Registration Guide

**TIME ALLOTTED:** 3 hours

**VERSION:** Currently testing on Alteryx Designer 20.4

**ATTEMPTS:** You must wait 2 months between your first and second attempt and 3 months between subsequent attempts

**EXAM TYPE:** Online, proctored, and performance-based

**QUESTIONS:** 19

**QUESTION TYPES:** Practical application scenarios, multiple-choice, and open form answer

**POLICY VIOLATIONS:** You may not receive assistance from another person, allow another person to take the exam on your behalf, plagiarize any exam answers, share exam content, create more than one account for additional exam attempts, or use unauthorized publications of exam answers. Review the Alteryx Certification Agreement and Policies for more details. Policy violations will result in removal of certifications and a possible ban from the Alteryx Community.

**RECERTIFICATION:** The Predictive Master certification expires after 2 years. You may retake and pass the Predictive Master exam OR any other Master-level exam OR the Expert exam to renew your Predictive Master Certification

# Performance-Based Exam Format

The exam is a performance-based assessment that allows you to prove your skill level and critical-thinking abilities in real-world scenarios. Certifications based on performance-based assessments are highly valuable because they test your ability to fully leverage a product to solve complex problems for an organization.

## Exam Format

The exam consists of seven real-word scenarios. In order to complete each scenario, you will perform a series of tasks in Designer. In addition to your Designer knowledge and skills, the scenarios test your critical-thinking and problem-solving abilities as you decide how to approach and solve each scenario. Instead of selecting a single answer, you demonstrate your solution to the scenario through the workflow and output files you create.

> The performance-based format of the exam allows you to do what you do best: create awesome solutions in Designer.

## Exam Delivery

The exam is delivered in a browser-based Windows virtual environment that has been configured especially for the exam. The environment has Alteryx Designer installed on it, as well as all the starting files and instructions needed to complete each scenario. The computer you use to take the exam needs to meet specific requirements to ensure the exam can be delivered properly. Refer to the User Guide for specific requirements.

# Exam Outline

The Alteryx Predictive Master Certification exam measures your ability to use the platform to solve complex predictive analytics problems. You should have a comprehensive knowledge of tools in the following palettes: Data Investigation, Predictive, Time Series, Predictive Grouping, and Prescriptive, and the ability to use them to perform complex tasks efficiently and effectively.

## Topic Distribution

- Classification – 28%
- Regression – 30%
- Clustering – 10%
- Time Series – 17%
- Optimization – 15%

## Prepare Data

- Impute missing values
- Handle unbalanced datasets
- Create training and validation datasets

## Investigate Data

- Perform feature engineering
- Reduce feature dimensionality using PCA
- Select predictor variables using correlation metrics
- Appropriately standardize predictor variables
- Interpret data investigation reports

## Train Models

- Train predictive models for classification analysis (Gradient Boosting, Logistic Regression, Naïve Bayes, Decision Tree, SVM, Random Forest, Neural Network)
- Train predictive models for regression analysis (Linear Regression, Gamma Regression, Linear regression with **stepwise** reduction of variables)
- Train models for time series forecasting (ETS and ARIMA)
- Train models for cluster analysis
- Train prescriptive models for optimization

## Compare Models

- Compare and evaluate predictive models using Model-Comparison and Cross-Validation tools
- Compare the performance of a full model vs. a model with reduced dimensionality
- Determine if two models are statistically equivalent
- Compare time series forecasting models

## Interpret Fit Statistics

- Interpret error measures (RMSE, MAPE, MASE)
- Compare precision measures of a class
- Compare F1 scores and average accuracy of models
- Interpret cluster information statistics

## Interpret and Apply Results

- Identify misclassified records
- Interpret coefficients of a linear regression model and determine the impact of specific variables
- Interpret a plot of residuals vs. fitted values
- Interpret confusion matrices
- Determine if an interaction effect exists
- Interpret PCA measures of proportion of variance
- Identify the best predictor variable
- Forecast future periods
- Score new data with a trained model
- Assign instances to clusters

# Exam Preparation

The best way to prepare for a performance-based assessment is through practical application of your skills. Nothing can replace the many hours you've spent working through complex scenarios in your professional work. These resources can help strengthen your skills, especially in areas you may not use regularly on the job.

## Data Science Learning Path

The field of Data Science covers a large number of areas that represent the future of analytics. With that in mind, we have created a Data Science Learning Path to help you get started. The Learning Path focuses on the R-based predictive tool palettes and includes interactive lessons covering:

- Predictive Modeling
- Time Series Forecasting
- Predictive Grouping
- Data Investigation concepts & strategies
- A rundown of some common predictive modeling algorithms and terminology

### Ways to Prepare:

Watch training videos and Interactive Lessons

Read blogs, articles, and documentation

Solve Weekly Challenges

Whether you're an aspiring Citizen Data Scientist or you just want to never stop learning, this course will help you get started! And don't forget to check out the Data Science blog & Tool Mastery series to supplement the Learning Path.

# Data Investigation

**WATCH:**

- Data Investigation Concepts (Data Science Learning Path)
- Data Investigation Techniques (Data Science Learning Path)

**READ:**

- How To: Complete Data Preparation and Investigation for Predictive Modeling
- Pre-Predictive: Using the Data Investigation Tools – Part 3
- Adventures in Data: Exploratory Data Analysis
- Tool Mastery: Association Analysis
- Tool Mastery: Test of Means

# Regression and Classification

**WATCH:**

- Predictive Modeling (Data Science Learning Path)
- Predictive Analytics Fundamentals (Data Science Learning Path)
- Creating a Predictive Model (Data Science Learning Path)

**READ**:

- Predictive Tools Flowchart
- What is a Confusion Matrix?
- Holdouts and Cross-Validation: Why the Data Used to Evaluate Your Model Matters
- Understanding the Outputs of the Decision Tree Tool
- An Overview of Stepwise Regression
- Bias Versus Variance
- Tool Mastery: Score Tool
- Model Comparison Tool
- Cross-Validation Tool

## TIME SERIES

**WATCH:**

- What is Time Series Forecasting? (Data Science Learning Path)
- Preparing Time Series Data (Data Science Learning Path)
- ETS and ARIMA (Data Science Learning Path)
- Selecting and Scaling Models (Data Science Learning Path)
- Time Series Modeling

**READ:**

- How to Use the ARIMA Tool
- How to Use the ETS Tool
- Champagne Analytics: A Time Series Tutorial
- Back to the Future: ARIMA and Forecasting with Covariates

**SOLVE:**

- 2019 Grand Prix US Heat 1
- When Will Community Hit 1000 Posts

## OPTIMIZATION

**WATCH:**

- Prescriptive Optimization
- Flex Your Prescriptive Optimization Muscles

**READ:**

- Tool Mastery: Optimization Tool
- Legolytics - Optimizing Cost
- Prescriptive Analytics: Unleash the Optimization Tool

**SOLVE:**

- Optimized Flower Arrangements

## CLUSTERING

**WATCH:**

- Predictive Grouping (Data Science Learning Path)
- Clustering in Designer (Data Science Learning Path)
- Principal Component Analysis (Data Science Learning Path)

**READ:**

- Tidying up with PCA: An Introduction to Principal Component Analysis
- Standardization in Cluster Analysis
- Tool Mastery: K-Centroids Cluster Analysis

## Additional Helpful Tips

The following more general information will also be useful for your exam preparation and during the exam itself. Read through these tips carefully!

- **Save your workflows frequently**! If you experience a technical issue on the exam but your workflow is saved, you will have a better chance of not having to repeat exam tasks.
- Make good use of the Browse Tool. Identify where you might want to view your model results before running your workflow.
- It's possible that some models may take a little time to run. It's good practice to use Tool Containers to disable some tools in order to have your workflow run faster and only on the models you choose.
- Consider saving your model results, either to a file or to an Input Data Tool. Just make sure to keep track of what you have saved and that you are creating your solution from the correct model results.

Begin with the scenario that you are most comfortable with. You can choose to work through the scenarios in any order, so start with one that you are most confident in your ability to solve.

# Appendix

**Alteryx Academy** – The Academy team provides Training content and programs to help the Alteryx Community users become accomplished and highly skilled Alteryx Designer users.

**Alteryx Certification Homepage** – The Certification page reflects the user's specific information; all available certifications will be listed as well as FAQ and Prep guides

**Alteryx Support** – Team within Alteryx that provides Standard, Advanced, and Premium levels of support to the Alteryx Community. Users are able to submit a case, read articles, and access helpful documentation for no cost.

**Certification** – Achievement earned by an individual to verify their legitimacy and competence to perform a job. A certification from Alteryx refers to a passing score on one of the full certification exams.

**Credly** – A well-known digital badge management software. It's an end-to-end solution for creating, issuing, and managing digital credentials. Every certification within our portfolio has a Credly badge associated with it to be shared virtually. Credly is also the home of our pdf certificates.

**Micro-Certification** – Achievement earned by an individual to verify their legitimacy and competence to perform a job. A micro-certification from Alteryx refers to a passing score on one of the smaller certification exams.

**Open Book** – Exam is not proctored, and test taker is allowed to use training, notes, and resources listed in this prep guide.

**Question types**: Multiple Choice – One Correct answer, Multiple Select – At least two correct answers and one incorrect answer, Matching, Practical Application - Questions in the practical application section gauge your ability to use the tools above to solve problems. Practical application questions require you to connect to a provided data set in Alteryx Designer to solve the problem.