

# Homework 5

*Group 1*

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Statement of the Problem</b>	<b>2</b>
<b>3</b>	<b>Data Exploration</b>	<b>2</b>
3.1	Variables Explained . . . . .	2

Prepared for:

Dr. Nathan Bastian

City University of New York, School of Professional Studies - Data 621

Prepared by:

Group 1

Senthil Dhanapal

Yadu Chittampalli

Christophe Hunt

# 1 Introduction

## 2 Statement of the Problem

## 3 Data Exploration

### 3.1 Variables Explained

The variables provided in the Wine Training Data Set are explained below:

Variable Code	Definition
<b>INDEX</b>	<b>Identification Variable (do not use)</b>
<b>TARGET</b>	Number of Cases Purchased
<b>AcidIndex</b>	<b>Proprietary method of testing total acidity of wine by using a weighted averag</b>
Alcohol	Alcohol Content
<b>Chlorides</b>	<b>Chloride content of wine</b>
CitricAcid	Citric Acid Content
<b>Density</b>	<b>Density of Wine</b>
FixedAcidity	Fixed Acidity of Wine
<b>FreeSulfurDioxide</b>	<b>Sulfur Dioxide content of wine</b>
LabelAppeal	Marketing Score indicating the appeal of label design for consumers. High numbers suggest customers like the label design. Negative numbers suggest customers don't like the design.
<b>ResidualSugar</b>	<b>Residual Sugar of wine</b>
STARS	Wine rating by a team of experts. 4 Stars = Excellent, 1 Star = Poor
<b>Sulphates</b>	<b>Sulfate content of wine</b>
TotalSulfurDioxide	Total Sulfur Dioxide of Wine
<b>VolatileAcidity</b>	<b>Volatile Acid content of wine</b>
pH	pH of wine