

**This assignment will be on Data Exploration**. The dataset needed for this are present on the server in the following location.

**Z:\ Foundation Exercises\** **Assignments\Class11 - Data Exploration and Preparation\Data Exploration**

1. Client XX is part of the horse racing industry in the US. They own and manage multiple race tracks across the US and Canada.

The VP Marketing was cold called by a sales person in your company pitching the insights that could be generated from data mining. He has sent you a limited data set and a data dictionary, and has asked that you demonstrate how you would generate insights of interest.

Your initial task is to assess the data. Use your knowledge of data exploration techniques to come up with an overall picture of the data.

Data - cdi\_model\_data.csv   
Data Dictionary is loaded in: CDI Data Dictionary

Use either Excel or SAS to assess the data. Save your findings in power point format. You can do this assignment outside of remote desktop. For sample datasets check with [support@jigsawacademy.com](mailto:support@jigsawacademy.com)

Useful links:

<http://en.wikipedia.org/wiki/Thoroughbred_horse_racing> tends to be European focused (which is different than US) but contains some good background information

More Wikipedia

links:<http://en.wikipedia.org/wiki/Horse_racing#Horse_racing_in_North_America>  [http://en.wikipedia.org/ wiki/Parimutuel\_gambling](http://en.wikipedia.org/wiki/Parimutuel_gambling)

National Thoroughbred Racing Association<http://www.ntra.com/default.aspx> Look specifically at the information under NTRA.U. There is a lot of good primer material here. The virtual Q&A is kind of cool.

While our websites are confusing to navigate, look here<http://www.twinspiresclub.com/playing_to_win/>for information on the basics of handicapping a race. Also look at websites for other major tracks as they typically have background information – Belmont,

Most popular industry

journals:<http://www.bloodhorse.com/><http://www.drf.com/><http://www.thoroughbredtimes.com/> [http:// www.horseplayerdaily.com/](http://www.horseplayerdaily.com/)<http://www.thoroughbreddailynews.com/>

Background on a Thoroughbred horse:<http://www.thejockeyclub.com/thoroughbredHistory.asp>Industry factoids:<http://www.thejockeyclub.com/factbook.asp>