Codebook for NASCAR dataset

There are 22 variables for each observation in the dataset. The first ten variables apply to the entire race, and so will be the same for all of the cars in a particular race. The remaining 12 variables apply to a specific driver.

Race Variables:

RaceNo

Race Number: sequential by year. The Daytona 500 is RaceNo 1 each year.

Date

The date of the race.

Site

The track at which the race took place.

Cars

The number of cars at the start of the race.

Len

The distance around the track.

Sfc

Track surface. For the Cup races included in this database, Sfc will always be either "P" for "Paved," or "R" for "Road Course." If the database is extended to include early years of NASCAR, some tracks may have surface "D" for "Dirt."

Miles

Lenth, in miles, of the race.

Cau

Number of Cautions during the race.

LC

Number of Lead Changes during the race.

Driver variables:

Driver

The name of the driver.

Finish

A numeric value indicating the order of finish of the driver.

Start

The ordinal position of the driver at the start of the race.

CarNumber

The number painted on the side of the car.

Make

The manufacturer of the car.

Pts

The number of championship points awarded to the driver, in the format (XX/Y). XX represents the total number of points, and the Y indicates the number of bonus points that go into that score. Since 2011, 43 points are awarded for 1st place, 42 for 2nd, down to 1 point for 43rd. The winning driver is awarded 3 bonus points for winning, and a driver can earn 1 bonus point for leading one or more laps, and 1 bonus point for leading the most laps. Therefore the highest possible score is 48/5, 48 total points of which 5 are bonus points. A different (and more complicated) system was used from 1975 to 2010.

Laps

The number of laps completed by that driver in that race.

Led

The number of laps led by that driver in that race.

Status

"Running" means that the car was still in competition at the end of the race. Other statuses indicate the reason the car was eliminated from the race before the end.

Rating

This statistic is only available in races from 2005 and later. Driver Rating is a Loop Data statistic which is based on several statistical components. A full explanation is in the appendix.

Winnings

The dollars awarded to the car/driver/team at the end of this race.

Team

The team which owns this car.

Apendix: Driver Rating Formula Explained

From NASCAR.com

Driver Rating is a Loop Data statistic which is based on several statistical components. Below is an explanation of the formula, which is broken down into three different sections.

SECTION 1 – Primary Statistics

The "Primary Statistics" are those that are most important to the formula – and therefore awarded the most points.

The points assigned to each "Primary Statistic" are based on the point structure in place when the Driver Rating was developed in 2005, and use the points structure in place from 2004-2006. In the interest of continuity, the formula still uses the old points structure, even though the point structure was changed in 2007.

In the formula, the first place driver earns 180 points and the 43rd-place driver earns 34, with the increments in between the same as NASCAR's points distribution scale. Each statistic is then "multiplied" or "weighted" depending on its importance to the formula. All ties in the "Primary Statistics" are broken by finishing position.

"Primary Statistic" points are then added together.

Primary Statistics	Max	Min	Multiplier
Finish	180	34	1
Average Running Positionˆ1	360	68	2
Average Speedˆ2	180	34	1
Fastest Lapˆ3	20	3.8	1/9

Footnotes:

ˆ1 Average Running Position are while on the lead lap and under a green flag.

ˆ2 For restrictor-plate races average speed in turns is used instead of overall average.

ˆ3 Fastest Lap is the average of the fastest three laps by that driver.

Example: At Darlington, Jimmie Johnson's Average Running Position ranked fourth. His point total would be 160, the point total of a driver who finished fourth in a race. In the Driver Rating formula, Average Running Position is important, and multiplied (or weighted) by two. Johnson's total for that particular Primary Statistic would be 320.

SECTION 2 – Fixed Bonus Points

Bonus points are given for reaching certain goals. Below are the goals and the bonus points for each. (Note: All ties in "fixed bonus points" result in the bonus being added to all applicable drivers.)

Each "Fixed Bonus Points" total is then added together.

Fixed Bonus Points	
Win	20
Top-15 Finish	10
Leading Most Laps	10
Lead Lap Finish	5
Average Running Position better than 10.0ˆ1	5
Average Running Position better than 6.0ˆ1	5
Average Running Position better than 2.0ˆ1	5

Footnotes:

ˆ1Average Running Position are while on the lead lap and under a green flag.

Example: At Darlington, Kyle Busch earned every "Fixed Bonus Points" except for Average Running Position better than 2.0. Therefore, in this section, he earned 55 points.

Example 2: At Darlington, both Jeff Gordon and Kyle Busch had an Average Running Position better than 6.0. They would each receive five points for having better than 10.0 and another five for having better than 6.0. Each would get a total of 10 for their Average Running Position.

SECTION 3 – Variable Bonus Points

These are bonus points that are dependent on two statistics a driver earns in a given race – green flag laps led and green flag fastest laps.

The two stats are added together and then divided by the total green flag laps the driver has run in the race. The resulting number is then multiplied by 100.

The maximum number of points a driver can get in this section is 100.

Example: At Darlington, Greg Biffle led 94 green flag laps and had 33 green flag fastest laps run. He ran 217 green flag laps in the race. The formula here is 94 + 33 = 127. Then: 127 / 217 = .585. Then: .585 x 100 = 58.5.

OVERVIEW

When all three sections are tallied, the points from each section are added together (the maximum number possible is 900 points) and then adjusted by dividing by six, which makes a perfect Driver Rating 150.0.

	Max	Min
Total Possible Points	900	139.8
Adjusted Driver Rating	150.0	23.3

Also, if a given race has less than 43 participants – the points awarded are adjusted accordingly to prevent inflated ratings. This is by design with reference to the NASCAR Craftsman Truck Series which has a standard field of 36 trucks.