Program Summary - 01 - CaseStudy17 - GettingTheData.sas

Execution Environment

Author:	sasdemo
File:	/folders/myshortcuts/myfolder/SSCode/20150724/01 - CaseStudy17 - GettingTheData.sas
SAS Platform:	Linux LIN X64 2.6.32-504.12.2.el6.x86_64
SAS Host:	LOCALHOST
SAS Version:	9.04.01M2P07232014
SAS Locale:	en_US
Submission Time:	7/25/2015, 12:44:26 PM
Browser Host:	10.0.2.2
User Agent:	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_3) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/44.0.2403.89 Safari/537.36
Application Server:	LOCALHOST

Code: 01 - CaseStudy17 - GettingTheData.sas

```
/*********************************************
/* FINAL CASE STUDY - HORSE RACING
/* Recommendations to a client on how to maximize handle
/* STEP 1: Get the Data and Filter the data to our analysis
/* ------------------------------------------------------
/* Customer is interested only in Tracks with IDs: CRC, AP, and FG
/* Also year of analysis should be limited to 2005 - 2006
/*********************************************/


/** Library names for the Original Dataset and the Working Datasets */
LIBNAME CS17 '/folders/myshortcuts/myfolder/Foundation Exercises/Assignments/Class17 - Final Case Study';
LIBNAME WK17 '/folders/myshortcuts/myfolder/SSCode/20150724/Datasets';


/** Creating A working TRACK Dataset filtering for Tracks with IDs:  CRC, AP, and FG */
DATA WK17.TRACK;
        SET CS17.track;
        WHERE
                track_id = "AP" OR
                track_id = "FG" OR
                track_id = "CRC";
RUN;

/** Creating a Working TRACK_ZONE Dataset **/
DATA WK17.TRACK_STATISTIC;
        SET CS17.track_statistic;
        WHERE
                (
                        track_id = "AP" OR
                        track_id = "FG" OR
                        track_id = "CRC"
                )  AND
                (
                        YEAR( DATEPART( race_date) ) = 2005 OR
                        YEAR( DATEPART( race_date) )= 2006
                );
        race_date_in_date = DATEPART( race_date);
        FORMAT race_date_in_date date9.;
RUN;


/** Creating a Working RACE Dataset **/
DATA WK17.RACE;
        SET CS17.RACE;
        WHERE
                (
                        track_id = "AP" OR
                        track_id = "FG" OR
                        track_id = "CRC"
                )  AND
                (
                        YEAR( DATEPART( race_date) ) = 2005 OR
                        YEAR( DATEPART( race_date) )= 2006
                );
        race_date_in_date = DATEPART( race_date);
        FORMAT race_date_in_date date9.;
RUN;


/** Creating a Working EXOTIC_PAYOFF dataset **/
DATA WK17.EXOTIC_PAYOFF;
        SET CS17.EXOTIC_PAYOFF;
        WHERE
                (
```

```
                        track_id = "AP" OR
                        track_id = "FG" OR
                        track_id = "CRC"
                )  AND
                (
                        YEAR( DATEPART( race_date) ) = 2005 OR
                        YEAR( DATEPART( race_date) )= 2006
                );
        race_date_in_date = DATEPART( race_date);
        FORMAT race_date_in_date date9.;
RUN;
```

Log: 01 - CaseStudy17 - GettingTheData.sas

Notes (18)

```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
57
58         /*************************************************
59         /* FINAL CASE STUDY - HORSE RACING
60         /* Recommendations to a client on how to maximize handle
61         /* STEP 1: Get the Data and Filter the data to our analysis
62         /* ------------------------------------------------------
63         /* Customer is interested only in Tracks with IDs: CRC, AP, and FG
64         /* Also year of analysis should be limited to 2005 - 2006
65         /************************************************/
66
67
68         /** Library names for the Original Dataset and the Working Datasets */
69         LIBNAME CS17 '/folders/myshortcuts/myfolder/Foundation Exercises/Assignments/Class17 - Final Case Study';
NOTE: Libref CS17 was successfully assigned as follows:
      Engine:        V9
      Physical Name: /folders/myshortcuts/myfolder/Foundation Exercises/Assignments/Class17 - Final Case Study
70         LIBNAME WK17 '/folders/myshortcuts/myfolder/SSCode/20150724/Datasets';
NOTE: Libref WK17 was successfully assigned as follows:
      Engine:        V9
      Physical Name: /folders/myshortcuts/myfolder/SSCode/20150724/Datasets
71
72
73         /** Creating A working TRACK Dataset filtering for Tracks with IDs:  CRC, AP, and FG */
74         DATA WK17.TRACK;
75         SET CS17.track;
NOTE: Data file CS17.TRACK.DATA is in a format that is native to another host, or the file encoding does not match the session
      encoding. Cross Environment Data Access will be used, which might require additional CPU resources and might reduce
      performance.
76         WHERE
77         track_id = "AP" OR
78         track_id = "FG" OR
79         track_id = "CRC";
80         RUN;

NOTE: There were 3 observations read from the data set CS17.TRACK.
      WHERE track_id in ('AP', 'CRC', 'FG');
NOTE: The data set WK17.TRACK has 3 observations and 12 variables.
NOTE: DATA statement used (Total process time):
      real time           0.16 seconds
      cpu time            0.08 seconds


81
82         /** Creating a Working TRACK_ZONE Dataset **/
83         DATA WK17.TRACK_STATISTIC;
84         SET CS17.track_statistic;
NOTE: Data file CS17.TRACK_STATISTIC.DATA is in a format that is native to another host, or the file encoding does not match t
      session encoding. Cross Environment Data Access will be used, which might require additional CPU resources and might red
      performance.
85         WHERE
86         (
87         track_id = "AP" OR
88         track_id = "FG" OR
89         track_id = "CRC"
90         )  AND
91         (
92         YEAR( DATEPART( race_date) ) = 2005 OR
93         YEAR( DATEPART( race_date) )= 2006
94         );
96         FORMAT race_date_in_date date9.;
97         RUN;

NOTE: DATA statement used (Total process time):
      real time           1.62 seconds
      cpu time            1.16 seconds

NOTE: There were 1585 observations read from the data set CS17.TRACK_STATISTIC.
      WHERE track_id in ('AP', 'CRC', 'FG') and YEAR(DATEPART(race_date)) in (2005, 2006);
NOTE: The data set WK17.TRACK_STATISTIC has 1585 observations and 12 variables.

98
99
100        /** Creating a Working RACE Dataset **/
101        DATA WK17.RACE;
```

```
102         SET CS17.RACE;
NOTE: Data file CS17.RACE.DATA is in a format that is native to another host, or the file encoding does not match the session
      encoding. Cross Environment Data Access will be used, which might require additional CPU resources and might reduce
      performance.
103         WHERE
104         (
105         track_id = "AP" OR
106         track_id = "FG" OR
107         track_id = "CRC"
108         )  AND
109         (
110         YEAR( DATEPART( race_date) ) = 2005 OR
111         YEAR( DATEPART( race_date) )= 2006
112         );
113         race_date_in_date = DATEPART( race_date);
114         FORMAT race_date_in_date date9.;
115         RUN;

NOTE: DATA statement used (Total process time):
      real time           4.84 seconds
      cpu time            3.67 seconds

NOTE: There were 6327 observations read from the data set CS17.RACE.
      WHERE track_id in ('AP', 'CRC', 'FG') and YEAR(DATEPART(race_date)) in (2005, 2006);
NOTE: The data set WK17.RACE has 6327 observations and 61 variables.

116
117
118         /** Creating a Working EXOTIC_PAYOFF dataset **/
119         DATA WK17.EXOTIC_PAYOFF;
120         SET CS17.EXOTIC_PAYOFF;
NOTE: Data file CS17.EXOTIC_PAYOFF.DATA is in a format that is native to another host, or the file encoding does not match the
      session encoding. Cross Environment Data Access will be used, which might require additional CPU resources and might red
      performance.
121         WHERE
122         (
123         track_id = "AP" OR
124         track_id = "FG" OR
125         track_id = "CRC"
126         )  AND
127         (
128         YEAR( DATEPART( race_date) ) = 2005 OR
129         YEAR( DATEPART( race_date) )= 2006
130         );
131         race_date_in_date = DATEPART( race_date);
132         FORMAT race_date_in_date date9.;
133         RUN;

NOTE: DATA statement used (Total process time):
      real time           6.73 seconds
      cpu time            4.96 seconds

NOTE: There were 26461 observations read from the data set CS17.EXOTIC_PAYOFF.
      WHERE track_id in ('AP', 'CRC', 'FG') and YEAR(DATEPART(race_date)) in (2005, 2006);
NOTE: The data set WK17.EXOTIC_PAYOFF has 26461 observations and 16 variables.

134
135
136
137
138         OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
150
```

Results: 01 - CaseStudy17 - GettingTheData.sas