

|  |  |  |  |
| --- | --- | --- | --- |
| Academic Year | 2022 | | |
| Semester | Fall | Winter | Summer |
| Course Code - Name | BAN130 | | |
| Instructor | Dr. Omar Altrad/Viji | | |
| Assessment | Projects |  | |
| Date | Tuesday, August 09, 2022 | | |

|  |  |  |
| --- | --- | --- |
| **Student ID** | **Student Name** | **Role** |
|  |  | Group Lead |
|  |  | Member |
|  |  | Member |
|  |  | Member |

Project deadline is midnight Tuesday, August 9, 2022

**Projects**

You are required to choose a project from the list of the projects specified in this document and complete it within groups of max. 4. The project is expected to test your technical skills in SAS base programming and soft skills.

Since, this is a group project, it is required to be done in **groups of minimum 3 and maximum 4.** In exceptional cases, there can be 1 member in a group. Each group should have a Group Lead who would be responsible for submitting the project on Blackboard (Please note that not all the members of the group are required to submit the project separately on Blackboard. One submission from the Group Lead would be sufficient).

The detailed requirements for each project are available in this document, so please go through the details and fulfil all the requirements to avoid missing any marks.

Finally, follow the below mentioned instructions carefully.

**Instructions:**

In order to obtain maximum marks in this assessment, please ensure the followings:

* Don’t forget to write your name and ID on the first page of this document. The student IDs and names of all the students in the group should be mentioned along with the roles.
* Submit the project by writing your solution in this document under the Solution heading below. Do not use a separate document. Everything related to the project should be included in this document, e.g., code, screenshots and etc.
* This project has a weightage of **30%** marks of the course.
* This is a group project so **only 1 submission from the group lead is required.**
* Group Leads are required to submit the project on Blackboard as instructed. Submissions through emails will not be accepted.
* The project deadline is **midnight Tuesday, August 9, 2022**. Submissions after the deadline will not be accepted.
* A separate session for presentation and QA for the project will be scheduled in week 12 or 13. Schedule will be uploaded on Blackboard.

**Rubric:**

Your assessment will be graded based on the following rubric:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Excellent (7 - 10)** | **Average (4 – 6.9)** | **Poor (<4)** |
| **Project Completion and Code**  **(10)** | The project was completed without any errors and output is as expected. Fulfills all/most of the requirements for the project. | The project was completed with few errors. Fulfills some of the requirements for the project. | The project is incomplete. Does not fulfill all/most of the requirements. |
| **Presentation and QA**  **(10)** | The student has a good contribution to the project. Knows ins and outs of the project.  The student has presented his/her part of the project very well. Knows everything / most of his/her part. | The student has average contribution to the project. Does not know the whole project.  The student has averagely presented his/her part of the project. Knows few of the things about his/her part. | The student has no contribution to the project. Does not know anything / most about the project.  The student has poorly presented the project. Does not know much about the project. |
| **Report**  **(10)** | Student has contributed well in preparing the project report and knows all the aspects of the report. | Student has contributed partially in preparing the project report and knows some aspects of the report. | Student has not contributed in preparing the report. |

**Project 1**

**Project Name:** Adventure Works Product Sales Analysis

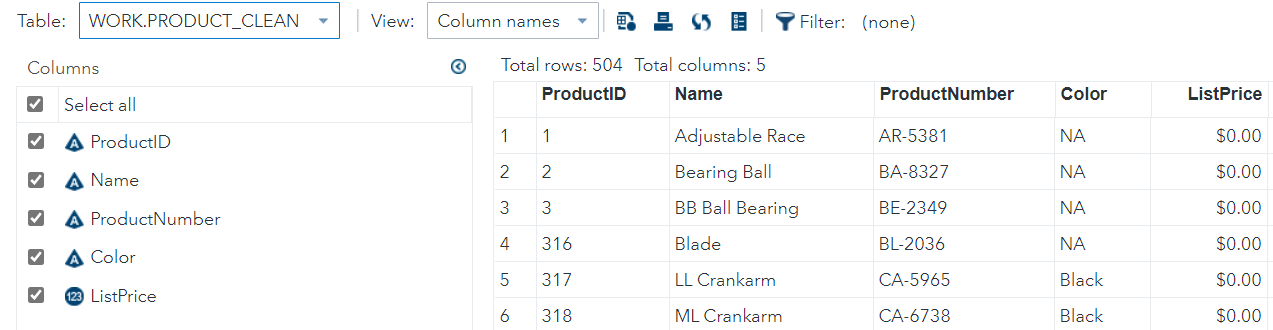
**Max. no. of students in a group:** 3-4 students

**Dataset:** AdventureWorks.xlsx (Available on Blackboard)

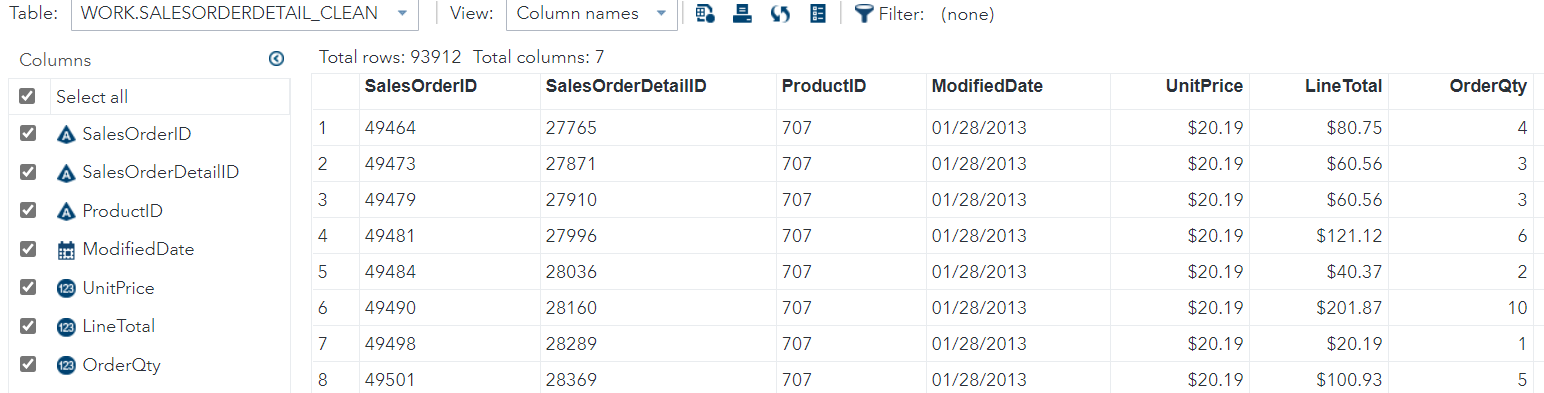
**Requirements:**

Below are bare minimum requirements for this project, however, you are free to add more features to your project:

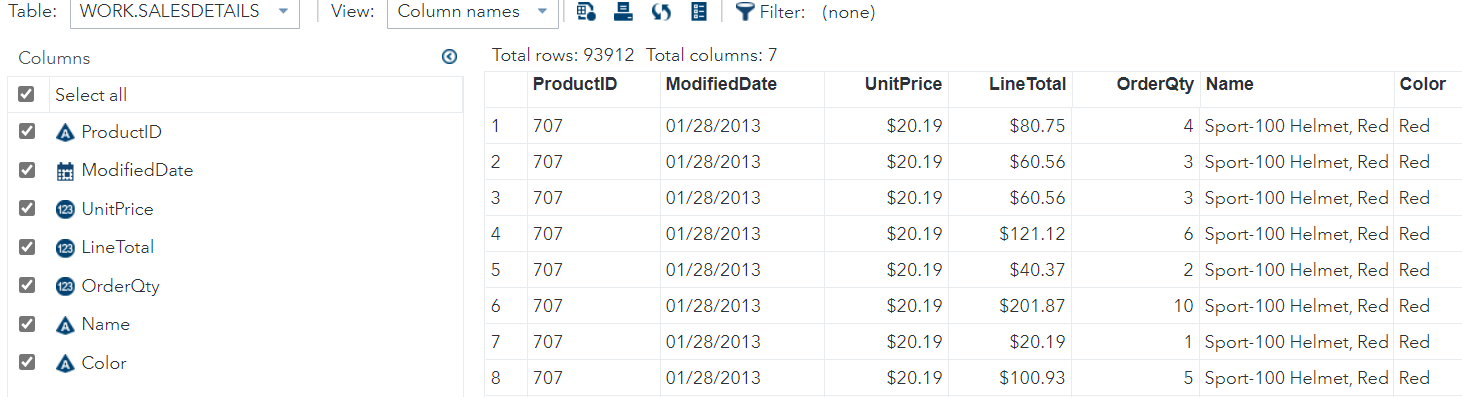
1. Data Import
   * This phase requires you to import the data from the provided excel file into SAS using Proc Import.
     + Product sheet in excel file should be imported as Product dataset in SAS.
     + SalesOrderDetail sheet in excel file should be imported as SalesOrderDetail dataset in SAS.
2. Data Cleaning
   * This phase requires you to clean your data before data analysis phase.
     + Product\_Clean:
       - Create a Product\_Clean dataset from Product dataset by bringing in only ProductID, Name, ProductNumber, Color and ListPrice
       - All the missing values in Color column should be replaced by ‘NA’
       - ListPrice column should be numeric (final column name should be ListPrice) and format should have a dollar sign with 2 decimal places
       - No un-necessary columns should be part of the Product\_Clean dataset. Please see below expected output.



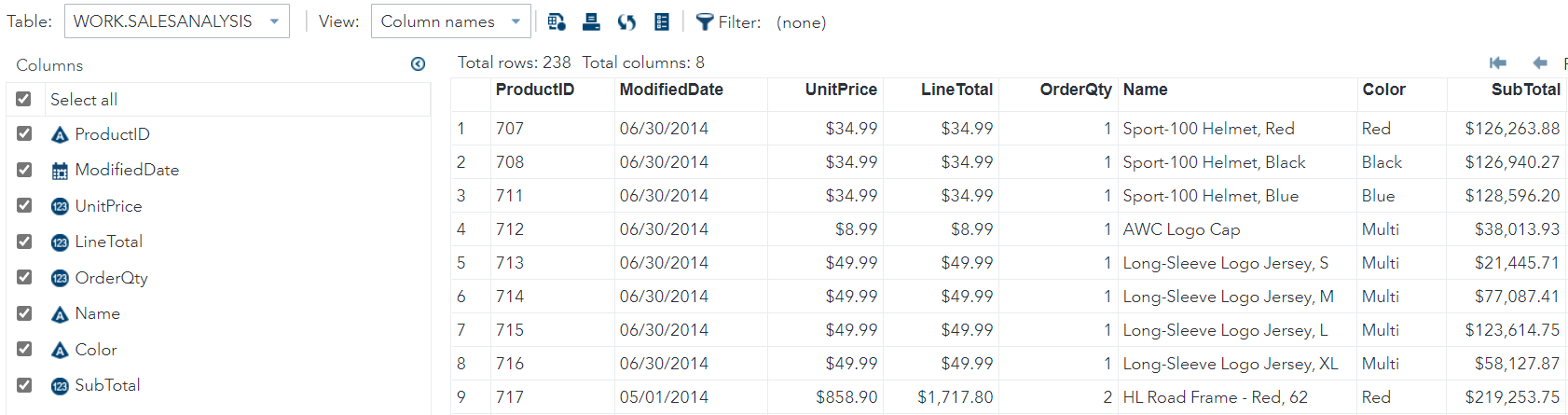
* + - SalesOrderDetail\_Clean:
      * Create SalesOrderDetail\_Clean dataset from SalesOrderDetail dataset by bringing in only SalesOrderID SalesOrderDetailID OrderQty ProductID UnitPrice LineTotal and ModifiedDate
      * ModifiedDate should be numeric with column name ModifiedDate
      * UnitPrice should be numeric with column name UnitPrice
      * LineTotal should be numeric with column name LineTotal
      * OrderQty should be numeric with column name OrderQty
      * Include date for year 2013 and 2014 in ModifiedDate only
      * ModifiedDate should be mmddyy10. Format
      * UnitPrice and LineTotal should have a dollar with 2 decimal places
      * No un-necessary columns should be part of the SalesOrderDetail\_Clean dataset. Please see expected output below:



1. Joining and Merging
   * This phase requires you to join / merge your datasets to create a dataset for analysis.
     + SalesDetails:
       - Create a SalesDetails dataset by joining SalesOrderDetail\_Clean and Product\_Clean datasets
       - Use ProductID column for joining the tables
       - SalesDetails table should contain all the observations from SalesOrderDetail\_Clean table along with columns from Product\_Clean
       - Drop SalesOrderID SalesOrderDetailID ProductNumber and ListPrice from the result dataset. Please see expected output below:



* + - SalesAnalysis:
      * Create a SalesAnalysis dataset from SalesDetails dataset that groups all the products by ProductID (hint: research on obtaining a total for each by group in SAS)
      * Create a SubTotal column in SalesAnalysis that provides an aggregate sum of each product by its ProductID.
      * SubTotal column should have.
      * Please see below expected output:



1. Data Analysis
   * This phase requires you to analyze the SalesAnalysis for Adventure Works and answer the following 5 questions by generating reports using Proc Print for each of the 5 questions:
     + How many Red color Helmets are sold in 2013 and 2014?
     + How many items sold in 2013 and 2014 have a Multi color?
     + What is the combined Sales total for all the helmets sold in 2013 and 2014?
     + How many Yellow Color Touring-1000 where sold in 2013 and 2014?
     + What was the total sales in 2013 and 2014?
   * Create at least one chart in SAS for any analysis of your choice from SalesAnalysis dataset (this analysis can be of your choice and not necessarily from above 5 questions.)
2. Project Report
   * This phase requires you to create a report in MS Word with the following requirements:
     + Explain each and every phase of the project (from Phase 1 to 4) along with the screenshots of the output and the related SAS code
     + Include answers to questions in Phase 4 in your report along with the chart you have chosen to create along with its justification
     + Make sure not to miss any phase and output of its screenshot

Phase 3:

SalesDetails:

Created a SalesDetails dataset by joining SalesOrderDetail\_Clean and Product\_Clean datasets

SalesAnalysis: Created a SalesAnalysis dataset from SalesDetails dataset that groups all the

products by ProductID.

**one-to-one join operation**

Primary key: ProductID For the merge to work, the first step is to sort both datasets

by the primary key.

IN= is used to generate temporary aliases for the input datasets. PROCESS: inner join-combining the two data sets and producing the data output where they intersect.

Does it appear in both the first and second tables? Otherwise, the data will not be sent to the final table.

The IF statement employs sub setting.

PROC SQL is a very powerful Base SAS7 Procedure that combines the functionality of DATA and PROC steps into a single step.

PROC SQL can sort, summaries, subset, join (merge), and concatenate datasets, create new variables, and print the results all in one step, as well as create a new table or view!

Calculated aggregate sum by product ID grouping.

The information is organised by ProductId. Salesanalysis table was created with the required variables from the salesdetails table. For the unique PRODUCT ID, the sum function was used for aggregate calculation, and it was named subtotal with a two-decimal format. Using group by SELECT, I grouped the data by productid. Proc sql statement: