

Team # _____

Topic	Grade Earned	Comments
Deliverables – Minus 20pts - All elements of the deliverables are submitted correctly (See Deliverables section in the Lab 3 document)		
Part #1 – 10pts		
- Final schema design Fact, Product Dim, Customer Dim, Order Date Dim, Sales Date Dim, Junk Dim (1pt each, total 7pts) - Rationale for final schema design (4pts)		
Part #2		
<i>Step 2-1: Extract Data & Formats (Appendix A – Data Cleansing) – 65pts</i> - PECinvoice.csv: <ul style="list-style-type: none"> ▪ Shifting columns (2pts) ▪ Incorrect date format (2pts) ▪ Incorrect shipping method (7pts) ▪ Sales amounts are incorrect (8pts) - PECproduct.csv: <ul style="list-style-type: none"> ▪ Unit costs are calculated using manufacturing costs (5pts) - PECcustomer.csv: <ul style="list-style-type: none"> ▪ Commas and periods in the address field (2pt) - TPCWproduct.csv: <ul style="list-style-type: none"> ▪ Separate supplier city and state into two fields: city and state (2pts) ▪ Duplication on product (5pts) - TPCWcustomer.csv: <ul style="list-style-type: none"> ▪ Commas and periods in the address field (2pt) 		

<ul style="list-style-type: none"> ▪ Department numbers/Suite number are included in the address field(2pt) ▪ Zip code with only four digits (2pt) <p>- TPCWinvoice.csv:</p> <ul style="list-style-type: none"> ▪ Shifting columns (2pt) ▪ One record with only invoiceID(2pts) ▪ Negative custID (2pt) ▪ Date in two different formats (2pt) ▪ Sales amounts are incorrect (2pt) <p>- Uniformity among three divisions (TPC-E, TPC-W, PEC)</p> <ul style="list-style-type: none"> ▪ Corp. verses Corporation (0.5pt) ▪ Co. verses Company (0.5pt) ▪ Street verses St. (0.5pt) ▪ Avenue verses Ave./Av. (0.5pt) ▪ Incorporated verses Inc/Inc. (0.5pt) ▪ Equip. verses equipment (0.5pt) <p>- Table 1, Appendix A is filled appropriately (4pts)</p> <p><i>Step 2-2: Agree on model – minus 6pts</i></p> <p>- Changes in the star schema are made according to the cleansing decisions made (minus 6pts if not)</p>		
Part #3		
<p><i>Step 3.1: Reformat the Data into “table images” for Loading – 40pts</i></p> <p>- Customer dimension</p> <ul style="list-style-type: none"> ▪ One format for customer type (1pts) ▪ One format for State field (1pts) ▪ Decisions on the address fields for customer (1pts) ▪ Duplications on customers (1pts) ▪ Create surrogate key for Customer dimension (1pt) <p>- Product dimension</p>		

<ul style="list-style-type: none"> ▪ Product type and business unit must be present for PEC, TPC-W, and TPC-E (3pts) ▪ Additional field (Sales by) that identifies the division who sales the product (1pts) ▪ Create surrogate key for Product dimension (1pt) <p><u>Note: a separate dimension can be created for Supplier but it must satisfy the below criteria.</u></p> <ul style="list-style-type: none"> ▪ Decisions on the address fields for supplier (1pts) ▪ One format for State field (1pts) ▪ TPC-W, TPC-E, and PEC must also be listed as suppliers since some of PEC products are made by PEC, TPC-W, or TPC-E (1pts) ▪ If separate Supplier dimension is created, duplications on supplier need to be taken care of (minus 2pts if not) ▪ Create surrogate key for Supplier dimension (minus 2pts if not) <ul style="list-style-type: none"> - Order Date dimension <ul style="list-style-type: none"> ▪ Fields in this dimension need to be generated (3pts) ▪ Create surrogate key for Order Date dimension (1pts) - Sales Date dimension <ul style="list-style-type: none"> ▪ Fields in this dimension need to be generated (3pts) ▪ Create surrogate key for Sales Date dimension (1pts) - Junk dimension <ul style="list-style-type: none"> ▪ Fields in this dimension need to be generated (3pts) ▪ Create surrogate key for Junk dimension (1pts) - Fact table 		
---	--	--

<ul style="list-style-type: none"> ▪ All the surrogate keys are presented (1pts) ▪ Null records in dimensions are created appropriately to handle non-match sales records (3pts) <p>- Table 2, Appendix A is filled appropriately (3pts)</p>		
Part #4 -30pts		
<p><i>Step 4.1: Load</i></p> <p>- Table 3, Appendix A is filled appropriately (1pts for each dimension loaded, 3 for fact; total 10pts)</p>		
Part #5 - 30pts		
<p><i>Step #3-5.1: Define three User Queries. For each user query</i></p> <ul style="list-style-type: none"> - Select one report goal (1pt) - Query to achieve the goal (5pts) - Appropriate index(es) (2pts) - Sample output (2pts) 		
TOTAL (out of 175):		