**SAS Competency Test**

**Part A:**

***18 Questions - Total Marks – 50 (Duration – 45 minutes)***

1. How SAS processes the data at the back end? **(2 Marks)**
2. What are the automatic variables created in PDV? And when their values get updated for the respective automatic variable? **(2 Marks)**
3. The below data is saved in a text file:

|  |  |  |
| --- | --- | --- |
| Name | Age | Gender |
| Amit | 18 |  |
| Kushal |  | M |
| Nakul | 20 | F |
| Sukhpreet | 27 | F |

Please write a SAS code to convert the text data (list data) into SAS data file. (The data values are separated with blanks) **(4 Marks)**

1. Difference between IF and WHERE statement. Which is more efficient and why? **(3 Marks)**
2. Can we calculate the sum of a Salary variable stored in WORK.XYZ dataset using data step? If so, please write an appropriate code to get the output. **(2 Marks)**
3. Write a sample code to get the cumulative SUM of salary variable in WORK.XYZ dataset. **(3 Marks)**
4. Difference between PROC Freq, PROC Means and PROC Summary. **(3 Marks)**
5. There are four variables (A, B, C & D) in PQR dataset. How can we get the cross-tabulation of A & D variables? **(2 Marks)**
6. There is a variables Age in SAS dataset i.e. “Demo”. Please create a new variable Age\_group to create different buckets for different age groups as shown below. Please do this task using data step as well as PROC SQL. **(4 Marks)**

|  |
| --- |
| Age |
| 18 |
| 19 |
| 26 |
| 32 |
| 45 |

|  |  |
| --- | --- |
| Age | Age\_Group |
| 18 | 0-18 |
| 19 | 19-25 |
| 26 | 26-35 |
| 32 | 26-35 |
| 45 | 35+ |

1. Difference between DATA merge and PROC SQL joins **(5 Marks)**
2. There is a SAS dataset i.e. “Demo”. We want to know about total IDs and unique IDs in this particular table. Please write a sample code to get this information. **(3 Marks)**
3. Difference between Keyword and Positional parameters in SAS Macros. **(2 Marks)**
4. There is a variables Age in SAS dataset i.e. “Demo”. We want to store all the Age values in different macro variables. Write a SAS code to store these values. **(5 Marks)**

|  |
| --- |
| Age |
| 18 |
| 19 |
| 26 |
| 32 |
| 45 |

1. For what purpose would you use the RETAIN statement? **(2 Marks)**
2. How do you control the number of observations read or written? **(2 Marks)**
3. Which function is used to remove the following blanks: **(2 Marks)**
   1. Leading
   2. Trailing
   3. Leading and trailing both
4. What are the different purposes of SUBSTR function? **(2 Marks)**
5. Purpose of ROUND function and its syntax. **(2 Marks)**

**Part B:**

***3 Questions - Total Marks – 25 (Duration – 1 Hr 15 minutes)***

**Instructions**

* Candidate needs to create the SAS datasets for the tables given in all the exercises below.

1. In the below tables, information about owners and their field information is given. Use DATA STEP to do the following:
   * 1. Generate a separate data set containing owner name, place, location, Yield1 and Yield2, and sum of Yield1 and Yield2 **(10 Marks)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Field Yield Data** | | | |  |  | **Ownership Data** | | |
| **Place** | **Location** | **Yield1** | **Yield2** |  |  | **Owner** | **Place** | **Location** |
| A | A | 2.95 | 51.30 |  |  | Hrshikesh | a | g |
| B | D | 39.17 | 8.09 |  |  | Teddy | d | k |
| C | F | 26.43 | 97.51 |  |  | Srinu | g | t |
| D | G | 56.53 | 64.05 |  |  | Rajpal | e | l |
| G | T | 52.25 | 35.33 |  |  | Raman | a | a |
| E | H | 80.43 | 66.39 |  |  | Kapil | a | n |
| A | G | 15.70 | 11.76 |  |  | Rakesh | b | d |
| B | F | 63.66 | 25.44 |  |  | Yuvi | c | f |
| C | R | 78.84 | 0.04 |  |  | Sibel | c | r |
| D | D | 12.75 | 0.14 |  |  | Yukta | d | t |
| G | T | 64.67 | 68.04 |  |  | Srinivas | d | g |
| E | Z | 41.20 | 73.02 |  |  | Dev | g | t |
| D | K | 5.17 | 23.51 |  |  | Vishwam | e | z |
| G | J | 92.16 | 11.43 |  |  | Mayank | b | m |
| E | L | 62.50 | 73.60 |  |  | Vivek | b | f |
| A | N | 3.17 | 84.54 |  |  | Vaibhav | c | r |
| B | M | 52.31 | 4.15 |  |  | Pawan | d | d |
| C | R | 39.22 | 16.89 |  |  | Mukul | g | j |
| D | T | 26.93 | 33.12 |  |  | Baljeet | e | H |

1. Names.sas file has 2 fields – name\_id and earning from different resources. Please note that each row corresponds to earning by person from a unique resource. Therefore a name\_id can have multiple earnings.
2. Write a query to store a list of unique name\_id in file unique.csv from names.sas **(3 Marks)**
3. Find out the average earning per person (name\_id) **(2 Marks)**
4. There are two files - File One contains information about id, name and File Two contains information about id, age and sex as shown in Table 1 and Table 2 below. Write a program to merge the two files to obtain the desired result as shown in Table 3 **(10 Marks)**

**Table1 FILE ONE**

|  |  |
| --- | --- |
| ID | NAME |
| A01 | SUE |
| A02 | TOM |
| A05 | KAY |
| A10 | JIM |

**Table2 FILE TWO**

|  |  |  |
| --- | --- | --- |
| ID | AGE | SEX |
| A01 | 58 | F |
| A02 | 20 | M |
| A04 | 47 | F |
| A10 | 11 | M |

**Table3 DESIRED OUTPUT**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | NAME | AGE | SEX |
| A01 | SUE | 58 | F |
| A02 | TOM | 20 | M |
| A04 |  | 47 | F |
| A10 | JIM | 11 | M |

**Table4 DESIRED OUTPUT**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | NAME | AGE | SEX |
| A01 | SUE | 58 | F |
| A02 | TOM | 20 | M |
| A10 | JIM | 11 | M |

**Table 5 DESIRED OUTPUT**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | NAME | AGE | SEX |
| A04 |  | 47 | F |

**Table 6 DESIRED OUTPUT**

|  |  |  |  |
| --- | --- | --- | --- |
| ID | NAME | AGE | SEX |
| A05 | KAY | . | . |