**Assignment 4- Transformation Commands on Similar Data**

**Objective**

**Simple Assignment: Transformation Commands on Similar Data**

This assignment will guide you in using Splunk’s transformation commands to extract, calculate, and analyze data from a similar dataset.

**Objective**

1. Ingest a .log file containing product sales data.
2. Use transformation commands (rex, eval, stats, table) to extract fields and perform calculations.
3. Generate meaningful insights and present the data in an organized format.

**Data Preparation**

Save the following as s

***ales\_data.log***

Laptop Electronics 5 1200

Smartphone Electronics 10 800

Shoes Fashion 20 50

Watch Fashion 15 200

Sofa Furniture 2 1000

Chair Furniture 8 150

**Steps to Complete the Assignment**

**1. Ingest the Data**

1. Navigate to **Settings > Add Data > Upload** in Splunk.
2. Upload the sales\_data.log file.
3. Assign the file to an index (e.g., sales\_data\_index) and a sourcetype (e.g., sales\_log).

**2. Queries and Tasks**

**Task 1: Extract Fields**

Extract fields Product, Category, Quantity, and Price using rex.

* **Query**:

index=sales\_data\_index sourcetype=sales\_log

| rex "(?<Product>\w+)\s(?<Category>\w+)\s(?<Quantity>\d+)\s(?<Price>\d+)"

| table Product, Category, Quantity, Price

* **Goal**: Display extracted fields in a table.

**Task 2: Calculate Total Revenue**

Calculate the total revenue for each product (Quantity \* Price).

* **Query**:

index=sales\_data\_index sourcetype=sales\_log

| rex "(?<Product>\w+)\s(?<Category>\w+)\s(?<Quantity>\d+)\s(?<Price>\d+)"

| eval total\_revenue=Quantity\*Price

| table Product, Quantity, Price, total\_revenue

* **Goal**: Add a new field total\_revenue and display it in a table.

**Task 3: Categorize Products by Price**

Add a Price Range field based on the Price:

* Price < 100: Low
* 100 ≤ Price < 500: Medium
* Price ≥ 500: High
* **Query**:

index=sales\_data\_index sourcetype=sales\_log

| rex "(?<Product>\w+)\s(?<Category>\w+)\s(?<Quantity>\d+)\s(?<Price>\d+)"

| eval price\_range=case(Price<100, "Low", Price<500, "Medium", Price>=500, "High")

| table Product, Price, price\_range

* **Goal**: Display products with their Price and categorized price\_range.

**Task 4: Count Products by Category**

Count the number of products in each category using stats.

* **Query**:

index=sales\_data\_index sourcetype=sales\_log

| rex "(?<Product>\w+)\s(?<Category>\w+)\s(?<Quantity>\d+)\s(?<Price>\d+)"

| stats count by Category

* **Goal**: Show the count of products grouped by category.

**Task 5: Find the Top 3 Highest Revenue Products**

Identify the top 3 products generating the most revenue.

* **Query**:

index=sales\_data\_index sourcetype=sales\_log

| rex "(?<Product>\w+)\s(?<Category>\w+)\s(?<Quantity>\d+)\s(?<Price>\d+)"

| eval total\_revenue=Quantity\*Price

| sort -total\_revenue

| head 3

| table Product, total\_revenue

* **Goal**: Display the top 3 products with their total\_revenue.

**Deliverables**

1. **Queries**:
   * Submit all SPL queries used in the assignment.
2. **Screenshots**:
   * Include screenshots of the results for each task.
3. **Insights**:
   * Write a short summary of findings, such as:
     + Total revenue for each product.
     + Distribution of products across price ranges.
     + Categories with the most products.
     + Top revenue-generating products.

**Expected Outcomes**

By completing this assignment, you will:

* Learn how to extract fields from raw log data.
* Use Splunk transformation commands to calculate and categorize data.
* Generate meaningful insights from sales data.