

```
customer_data = customer_data.dropna(subset=['TotalValue', 'Quantity'])
```

```
customer_data_scaled = customer_data[['TotalValue', 'Quantity']]
```

```
customer_data_scaled = (customer_data_scaled - customer_data_scaled.mean()) /  
customer_data_scaled.std()
```

```
similarity_matrix = cosine_similarity(customer_data_scaled)
```

```
import numpy as np
```

```
def get_top_3_lookalikes(customer_id, similarity_matrix, customer_ids):
```

```
    customer_ids_list = customer_ids.tolist()
```

```
    idx = customer_ids_list.index(customer_id)
```

```
    sim_scores = similarity_matrix[idx]
```

```
    top_3_indices = np.argsort(sim_scores)[-4:-1]
```

```
    return [(customer_ids_list[i], sim_scores[i]) for i in top_3_indices]
```

```
lookalikes = {}
```

```
for customer_id in customer_data['CustomerID'][:20]:
```

```
    top_3 = get_top_3_lookalikes(customer_id, similarity_matrix, customer_data['CustomerID'].values)
```

```
    lookalikes[customer_id] = top_3
```

```
lookalike_df = pd.DataFrame([(key, val[0][0], val[0][1], val[1][0], val[1][1], val[2][0], val[2][1])
```

```
                             for key, val in lookalikes.items()] ,
```

```
                           columns=['CustomerID', 'Lookalike1_ID', 'Lookalike1_Score', 'Lookalike2_ID',  
'Lookalike2_Score', 'Lookalike3_ID', 'Lookalike3_Score'])
```

```
lookalike_df.to_csv('Lookalike.csv', index=False)
```