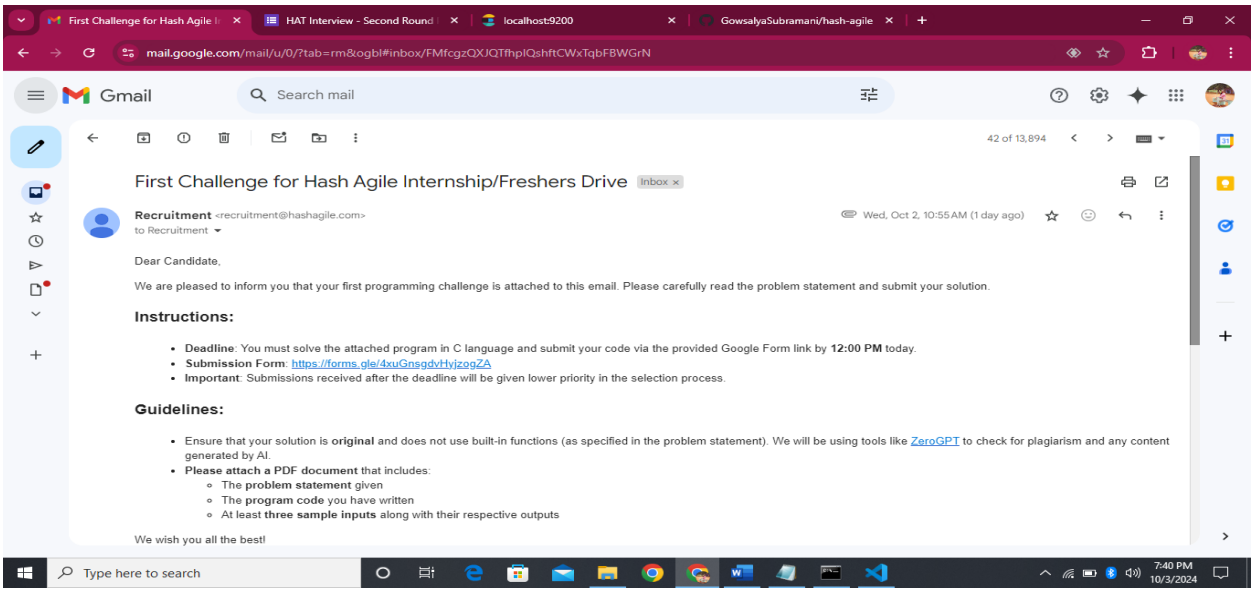
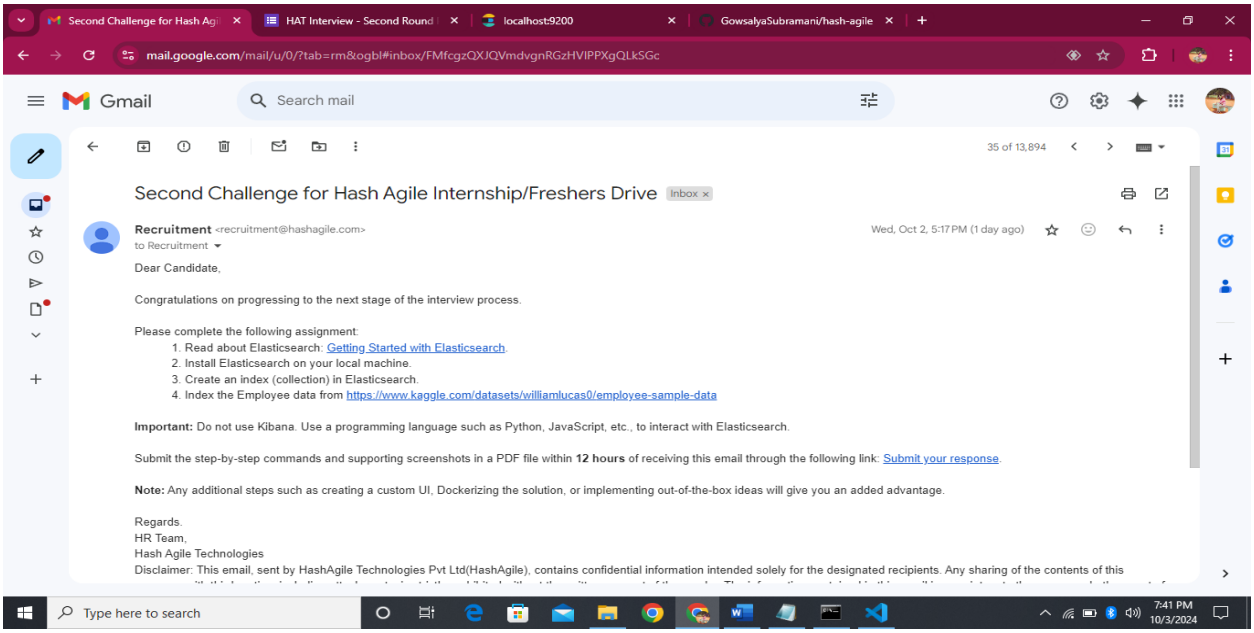
INTERVIEW ASSIGNMENT TASK

1.FULL NAME: GOWSALYA S

2. First Task Email Screenshot:



3.Second Task Email Screenshot:



4.&5. Github URL for Round 1 and Assignment

https://github.com/GowsalyaSubramani/hash-agile.git

6. Function Execution Results:

elasticsearch_employee.py

```
import csv
from elasticsearch import Elasticsearch
from elasticsearch.exceptions import NotFoundError

# Initialize Elasticsearch client
es = Elasticsearch(
    [{'scheme': 'http', 'host': 'localhost', 'port': 9200}],
    basic_auth=('elastic', 's_A0pGy+sJY-Z5OvyO8-') ,
     timeout = 30 # Replace 'your_password' with the actual password
)
try:
    es.ping()
    print("Successfully connected to Elasticsearch!")
except Exception as e:
    print(f"Connection failed: {e}")

# Function to create a collection (index)
def createCollection(p_collection_name):
    if not es.indices.exists(index=p_collection_name):
        es.indices.create(index=p_collection_name)
        print(f"Index '{p_collection_name}' created.")
```

```
    else:
        print(f"Index '{p_collection_name}' already exists.")

# Function to index data
def indexData(p_collection_name, p_exclude_column):
    with open(r'C:\Users\Hi\Downloads\Employee Sample Data 1.csv', newline='') as file:
        reader = csv.DictReader(file)
        for i, row in enumerate(reader):
            if p_exclude_column in row:
                del row[p_exclude_column]
            es.index(index=p_collection_name, id=i+1, document=row)
    print(f"Data indexed into '{p_collection_name}', excluding column '{p_exclude_column}'.")

# Function to search by column
def searchByColumn(p_collection_name, p_column_name, p_column_value):
    query = {
        "query": {
            "match": {
                p_column_name: p_column_value
            }
        }
    }
    response = es.search(index=p_collection_name, body=query)
    for hit in response['hits']['hits']:
        print(hit['_source'])

# Function to get employee count
def getEmpCount(p_collection_name):
```

```
    count = es.count(index=p_collection_name)['count']
    print(f"Total employees in '{p_collection_name}': {count}")
    return count

# Function to list
def list_documents(p_collection_name):
    response = es.search(index=p_collection_name, body={"query": {"match_all": {}}})
    documents = response['hits']['hits']
    if not documents:
        print(f"No documents found in index '{p_collection_name}'.")
    else:
        print(f"Documents in index '{p_collection_name}':")
        for doc in documents:
            print(f"ID: {doc['_id']}, Source: {doc['_source']}")

# Function to get documentID
def getDocumentIdByEmployeeId(p_collection_name, employee_id):
    response = es.search(index=p_collection_name, body={
        "query": {
            "match": {
                "Employee ID": employee_id
            }
        }
    })
    if response['hits']['total']['value'] > 0:
        for hit in response['hits']['hits']:
            print(f"Found Document ID: {hit['_id']} for Employee ID: {employee_id}")
            return hit['_id']  # Return the Elasticsearch document ID
```

```
    else:
        print(f"No employee found with Employee ID {employee_id}")
        return None

# Function to delete employee by ID
def delEmpById(p_collection_name, p_employee_id):
    try:
        es.delete(index=p_collection_name, id=p_employee_id)
        print(f"Employee with ID '{p_employee_id}' deleted from '{p_collection_name}'.")
    except NotFoundError as e:
        print(f"Error: Employee with ID '{p_employee_id}' not found in '{p_collection_name}'.")
    except Exception as e:
        print(f"An error occurred: {e}")

# Function to get department facet
def getDepFacet(p_collection_name):
    query = {
        "size": 0,
        "aggs": {
            "group_by_department": {
                "terms": {
                    "field": "Department.keyword"
                }
            }
        }
    }
    response = es.search(index=p_collection_name, body=query)
    buckets = response['aggregations']['group_by_department']['buckets']
```

```
    for bucket in buckets:
        print(f"Department: {bucket['key']}, Count: {bucket['doc_count']}")

# Function execution
v_nameCollection = 'hash_gowsalya'  # Use lowercase
v_phoneCollection = 'hash_1234'      # Replace '1234' with your last four digits

createCollection(v_nameCollection)
createCollection(v_phoneCollection)

getEmpCount(v_nameCollection)

indexData(v_nameCollection, 'Department')
indexData(v_phoneCollection, 'Gender')

# Call this function to list documents before trying to delete
list_documents(v_nameCollection)

# Change this to the actual Employee ID you want to delete
employee_id_to_delete = 'E02591'  # Employee ID you want to delete
doc_id = getDocumentIdByEmployeeId(v_nameCollection, employee_id_to_delete)

if doc_id:
    delEmpById(v_nameCollection, doc_id)  # Use the Elasticsearch document ID for deletion

delEmpById(v_nameCollection, 'E02591')

getEmpCount(v_nameCollection)
```

searchByColumn(v_nameCollection, 'Department', 'IT')

searchByColumn(v_nameCollection, 'Gender', 'Male')

searchByColumn(v_phoneCollection, 'Department', 'IT')

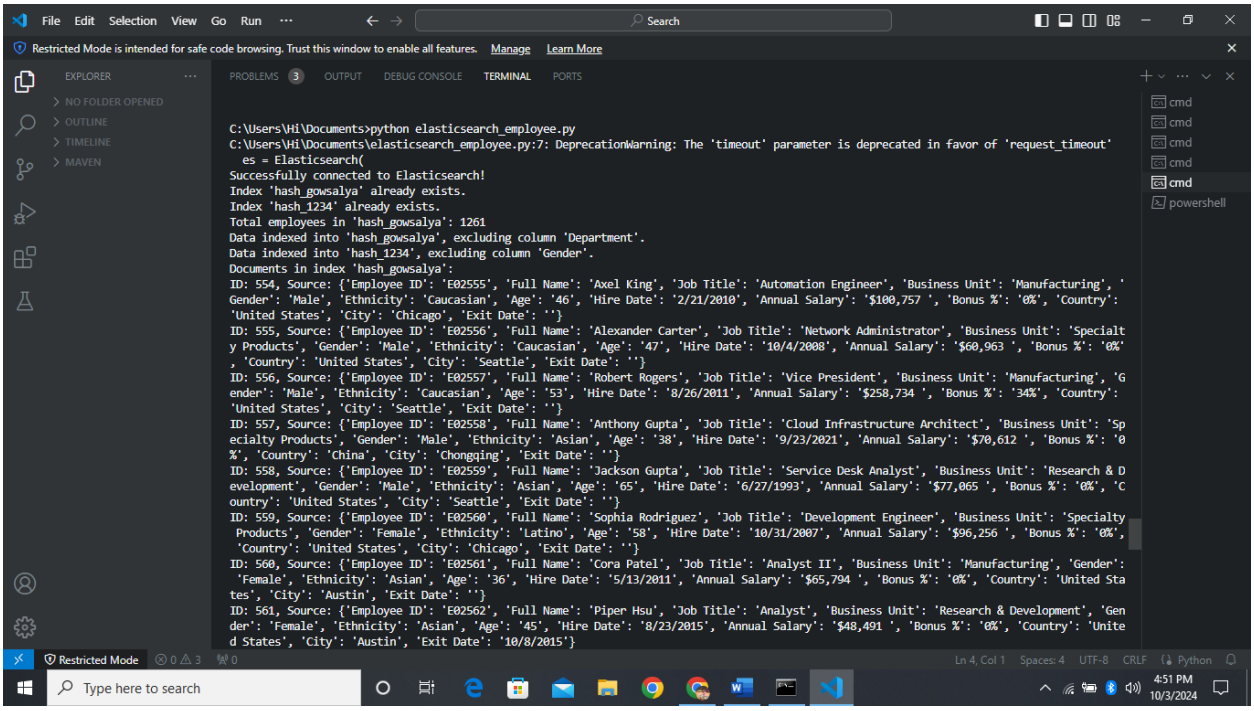
getDepFacet(v_nameCollection)

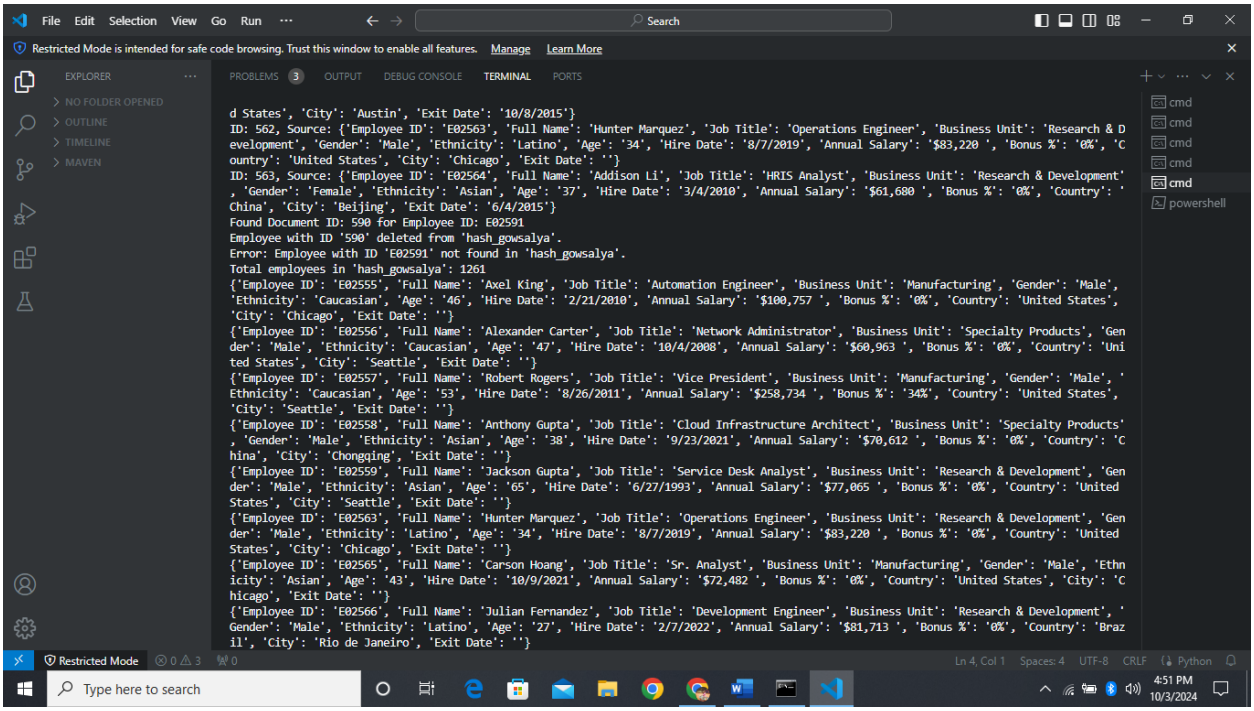
getDepFacet(v_phoneCollection)

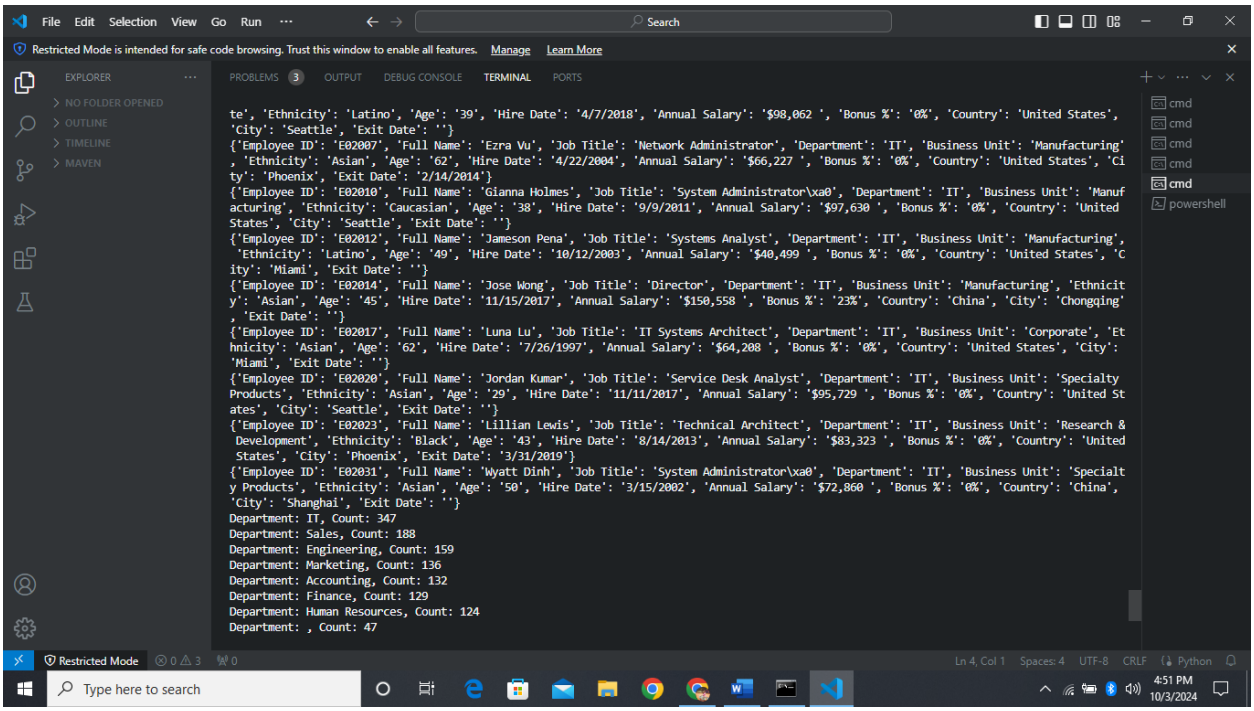
OUTPUT:

Run this command to see the below output:

python elasticsearch_employee.py




File Edit Selection View Go Run
Search
Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More
EXPLORER
NO FOLDER OPENED
OUTLINE
TIMELINE
MAVEN
PROBLEMS 3 OUTPUT DEBUG CONSOLE TERMINAL PORTS
d States', 'City': 'Austin', 'Exit Date': '10/8/2015'}
ID: 562, Source: {'Employee ID': 'E02563', 'Full Name': 'Hunter Marquez', 'Job Title': 'Operations Engineer', 'Business Unit': 'Research & Development', 'Gender': 'Male', 'Ethnicity': 'Latino', 'Age': '34', 'Hire Date': '8/7/2019', 'Annual Salary': '$83,220 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Chicago', 'Exit Date': ''}
ID: 563, Source: {'Employee ID': 'E02564', 'Full Name': 'Addison Li', 'Job Title': 'HRIS Analyst', 'Business Unit': 'Research & Development', 'Gender': 'Female', 'Ethnicity': 'Asian', 'Age': '37', 'Hire Date': '3/4/2010', 'Annual Salary': '$61,680 ', 'Bonus %': '0%', 'Country': 'China', 'City': 'Beijing', 'Exit Date': '6/4/2015'}
Found Document ID: 590 for Employee ID: E02591
Employee with ID '590' deleted from 'hash_gowsalya'.
Error: Employee with ID 'E02591' not found in 'hash_gowsalya'.
Total employees in 'hash_gowsalya': 1261
{'Employee ID': 'E02555', 'Full Name': 'Axel King', 'Job Title': 'Automation Engineer', 'Business Unit': 'Manufacturing', 'Gender': 'Male', 'Ethnicity': 'Caucasian', 'Age': '46', 'Hire Date': '2/21/2010', 'Annual Salary': '$100,757 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Chicago', 'Exit Date': ''}
{'Employee ID': 'E02556', 'Full Name': 'Alexander Carter', 'Job Title': 'Network Administrator', 'Business Unit': 'Specialty Products', 'Gender': 'Male', 'Ethnicity': 'Caucasian', 'Age': '47', 'Hire Date': '10/4/2008', 'Annual Salary': '$60,963 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02557', 'Full Name': 'Robert Rogers', 'Job Title': 'Vice President', 'Business Unit': 'Manufacturing', 'Gender': 'Male', 'Ethnicity': 'Caucasian', 'Age': '53', 'Hire Date': '8/26/2011', 'Annual Salary': '$258,734 ', 'Bonus %': '34%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02558', 'Full Name': 'Anthony Gupta', 'Job Title': 'Cloud Infrastructure Architect', 'Business Unit': 'Specialty Products', 'Gender': 'Male', 'Ethnicity': 'Asian', 'Age': '38', 'Hire Date': '9/23/2021', 'Annual Salary': '$70,612 ', 'Bonus %': '0%', 'Country': 'China', 'City': 'Chongqing', 'Exit Date': ''}
{'Employee ID': 'E02559', 'Full Name': 'Jackson Gupta', 'Job Title': 'Service Desk Analyst', 'Business Unit': 'Research & Development', 'Gender': 'Male', 'Ethnicity': 'Asian', 'Age': '65', 'Hire Date': '6/27/1993', 'Annual Salary': '$77,065 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02563', 'Full Name': 'Hunter Marquez', 'Job Title': 'Operations Engineer', 'Business Unit': 'Research & Development', 'Gender': 'Male', 'Ethnicity': 'Latino', 'Age': '34', 'Hire Date': '8/7/2019', 'Annual Salary': '$83,220 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Chicago', 'Exit Date': ''}
{'Employee ID': 'E02565', 'Full Name': 'Carson Hoang', 'Job Title': 'Sr. Analyst', 'Business Unit': 'Manufacturing', 'Gender': 'Male', 'Ethnicity': 'Asian', 'Age': '43', 'Hire Date': '10/9/2021', 'Annual Salary': '$72,482 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Chicago', 'Exit Date': ''}
{'Employee ID': 'E02566', 'Full Name': 'Julian Fernandez', 'Job Title': 'Development Engineer', 'Business Unit': 'Research & Development', 'Gender': 'Male', 'Ethnicity': 'Latino', 'Age': '27', 'Hire Date': '2/7/2022', 'Annual Salary': '$81,713 ', 'Bonus %': '0%', 'Country': 'Brazil', 'City': 'Rio de Janeiro', 'Exit Date': ''}
cmd
cmd
cmd
cmd
cmd
powershell
Restricted Mode
Ln 4, Col 1 Spaces: 4 UTF-8 CRLF Python
Type here to search
4:51 PM
10/3/2024


File Edit Selection View Go Run
Search
Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More
EXPLORER
NO FOLDER OPENED
OUTLINE
TIMELINE
MAVEN
PROBLEMS 3 OUTPUT DEBUG CONSOLE TERMINAL PORTS
te', 'Ethnicity': 'Latino', 'Age': '39', 'Hire Date': '4/7/2018', 'Annual Salary': '$98,062 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02007', 'Full Name': 'Ezra Vu', 'Job Title': 'Network Administrator', 'Department': 'IT', 'Business Unit': 'Manufacturing', 'Ethnicity': 'Asian', 'Age': '62', 'Hire Date': '4/22/2004', 'Annual Salary': '$66,227 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Phoenix', 'Exit Date': '2/14/2014'}
{'Employee ID': 'E02010', 'Full Name': 'Gianna Holmes', 'Job Title': 'System Administrator\xa0', 'Department': 'IT', 'Business Unit': 'Manufacturing', 'Ethnicity': 'Caucasian', 'Age': '38', 'Hire Date': '9/9/2011', 'Annual Salary': '$97,630 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02012', 'Full Name': 'Jameson Pena', 'Job Title': 'Systems Analyst', 'Department': 'IT', 'Business Unit': 'Manufacturing', 'Ethnicity': 'Latino', 'Age': '49', 'Hire Date': '10/12/2003', 'Annual Salary': '$40,499 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Miami', 'Exit Date': ''}
{'Employee ID': 'E02014', 'Full Name': 'Jose Wong', 'Job Title': 'Director', 'Department': 'IT', 'Business Unit': 'Manufacturing', 'Ethnicity': 'Asian', 'Age': '45', 'Hire Date': '11/15/2017', 'Annual Salary': '$150,558 ', 'Bonus %': '23%', 'Country': 'China', 'City': 'Chongqing', 'Exit Date': ''}
{'Employee ID': 'E02017', 'Full Name': 'Luna Lu', 'Job Title': 'IT Systems Architect', 'Department': 'IT', 'Business Unit': 'Corporate', 'Ethnicity': 'Asian', 'Age': '62', 'Hire Date': '7/26/1997', 'Annual Salary': '$64,208 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Miami', 'Exit Date': ''}
{'Employee ID': 'E02020', 'Full Name': 'Jordan Kumar', 'Job Title': 'Service Desk Analyst', 'Department': 'IT', 'Business Unit': 'Specialty Products', 'Ethnicity': 'Asian', 'Age': '29', 'Hire Date': '11/11/2017', 'Annual Salary': '$95,729 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Seattle', 'Exit Date': ''}
{'Employee ID': 'E02023', 'Full Name': 'Lillian Lewis', 'Job Title': 'Technical Architect', 'Department': 'IT', 'Business Unit': 'Research & Development', 'Ethnicity': 'Black', 'Age': '43', 'Hire Date': '8/14/2013', 'Annual Salary': '$83,323 ', 'Bonus %': '0%', 'Country': 'United States', 'City': 'Phoenix', 'Exit Date': '3/31/2019'}
{'Employee ID': 'E02031', 'Full Name': 'Wyatt Dinh', 'Job Title': 'System Administrator\xa0', 'Department': 'IT', 'Business Unit': 'Specialty Products', 'Ethnicity': 'Asian', 'Age': '50', 'Hire Date': '3/15/2002', 'Annual Salary': '$72,860 ', 'Bonus %': '0%', 'Country': 'China', 'City': 'Shanghai', 'Exit Date': ''}
Department: IT, Count: 347
Department: Sales, Count: 188
Department: Engineering, Count: 159
Department: Marketing, Count: 136
Department: Accounting, Count: 132
Department: Finance, Count: 129
Department: Human Resources, Count: 124
Department: , Count: 47
cmd
cmd
cmd
cmd
cmd
powershell
Restricted Mode
Ln 4, Col 1 Spaces: 4 UTF-8 CRLF Python
Type here to search
4:51 PM
10/3/2024