

**TUGAS 5**  
**METODE PENGUMPULAN DATA DIGITAL (B)**  
**WEB SCRAPING API DENGAN VARIABEL**



**Disusun Oleh:**

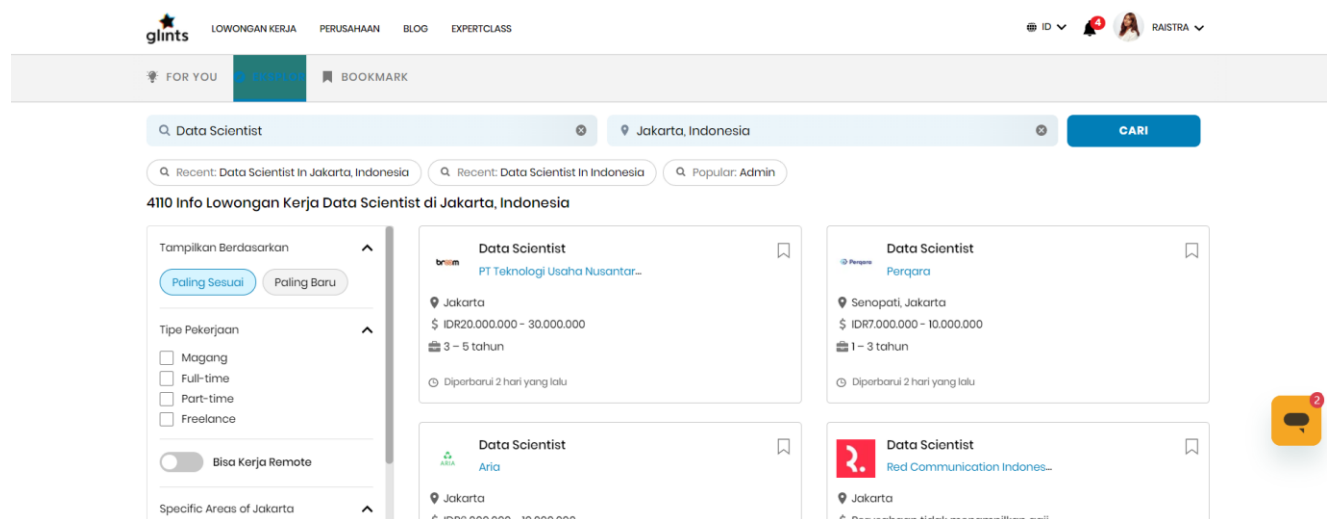
Raistra Aurel

2006572876

**PROGRAM STUDI STATISTIKA**  
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**  
**UNIVERSITAS INDONESIA**  
**2023**

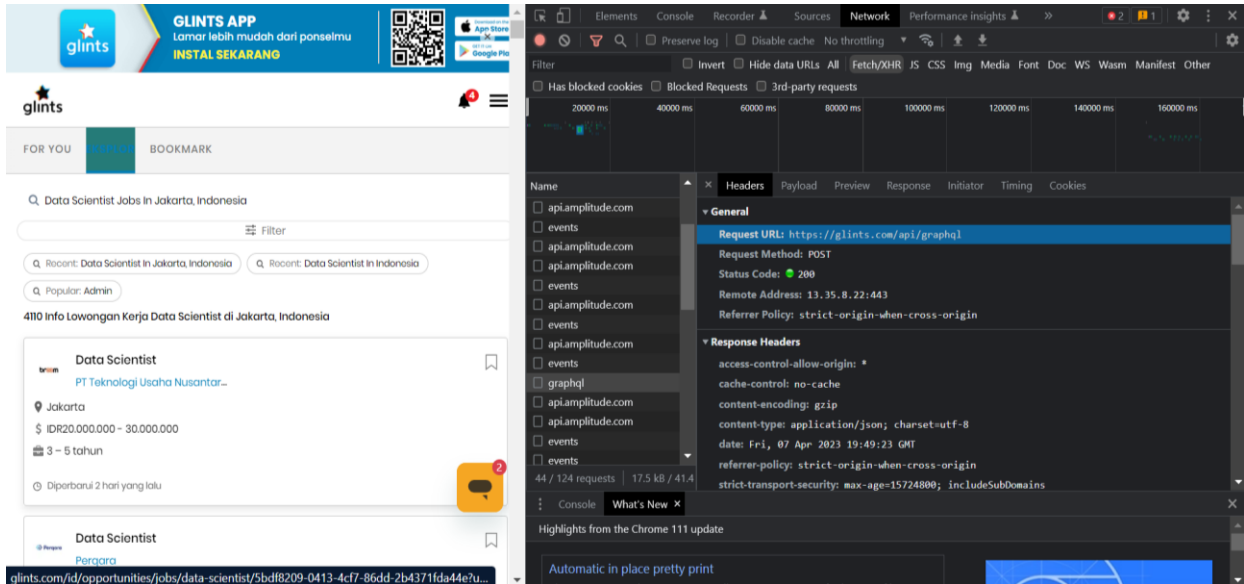
## 1. Website

Website yang akan digunakan untuk melakukan web scraping adalah website Glints (<https://glints.com/id>) Glints merupakan platform online yang berfokus pada rekrutmen dan pengembangan karir di Asia Tenggara. Melalui Glints, pengguna dapat mencari dan melamar pekerjaan, serta membangun jaringan profesional sesama pengguna dan perusahaan. Pencarian spesifik yang digunakan adalah untuk pekerjaan “Data Scientist” lokasi Jakarta, Indonesia pada link berikut <https://glints.com/id/opportunities/jobs/explore?keyword=Data+Scientist&country=ID&searchCity=28904&locationName=Jakarta%2C+Indonesia>



## 2. API

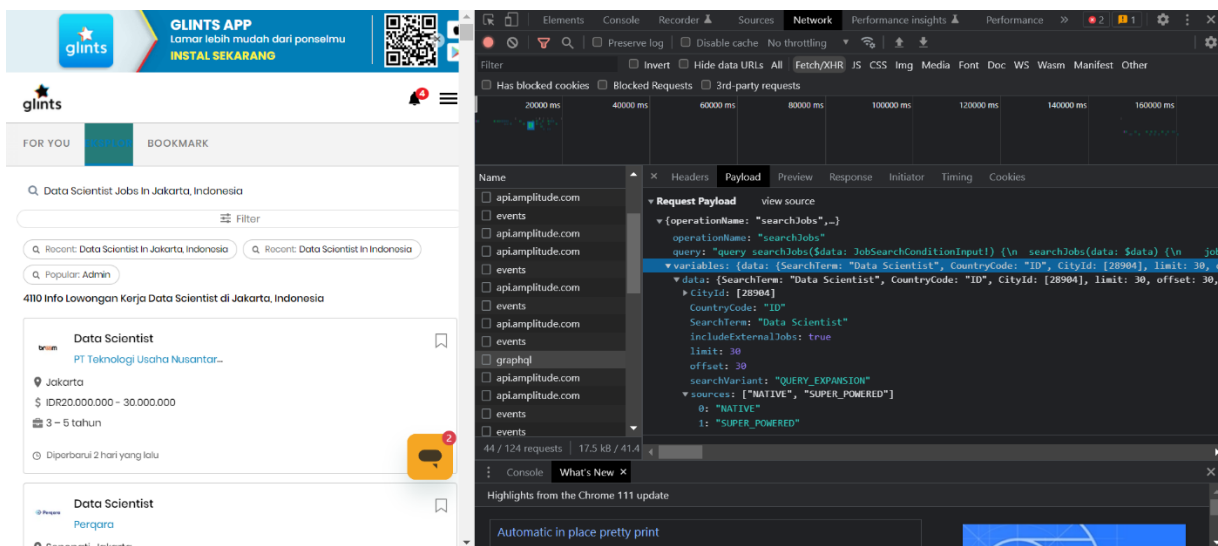
Pencarian API dilakukan dengan inspect (*f12 key*) pada situs. API yang dicari adalah API dengan variable dengan request method POST serta Status Code 200, ditemukan api dengan url berikut: <https://glints.com/api/graphql>



### 3. Variabel

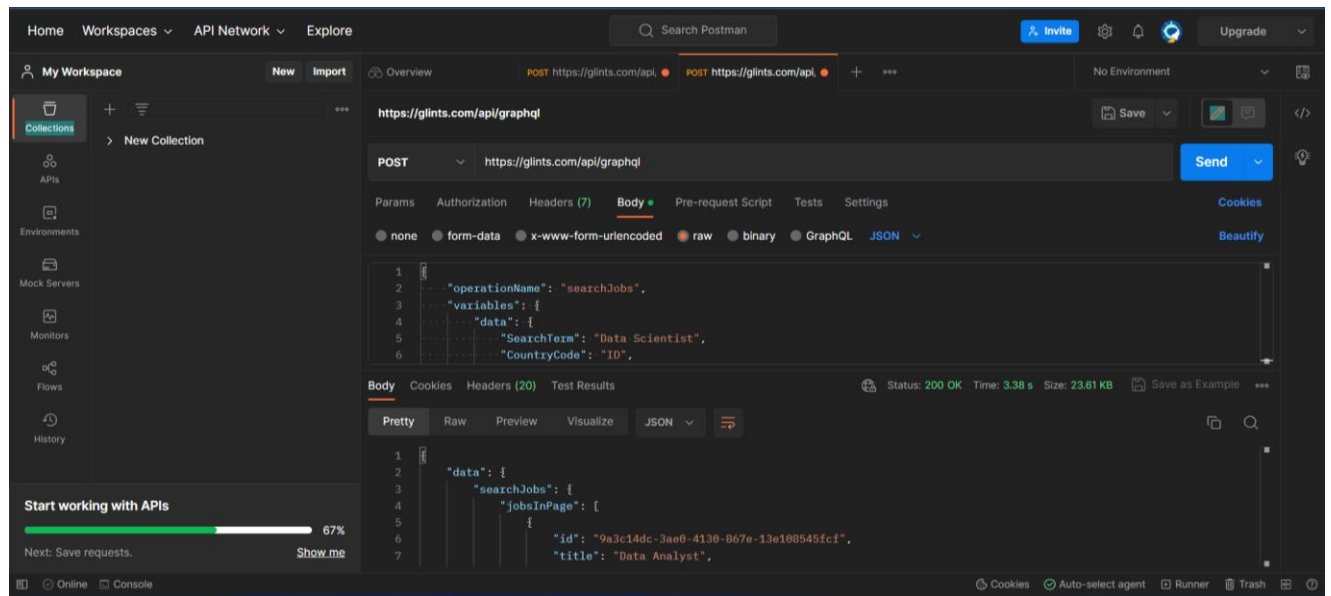
Dapat dilihat variabel pada *Payload*, terdapat variabel:

1. CityID: [28904]
2. CountryCode: "ID"
3. SearchTerm: "Data Scientist"
4. includeExternalJobs: true
5. limit: 30
6. offset: 30
7. searchVariant: "QUERY\_EXPANSION"
8. sources: ["NATIVE", "SUPER\_POWERED"]

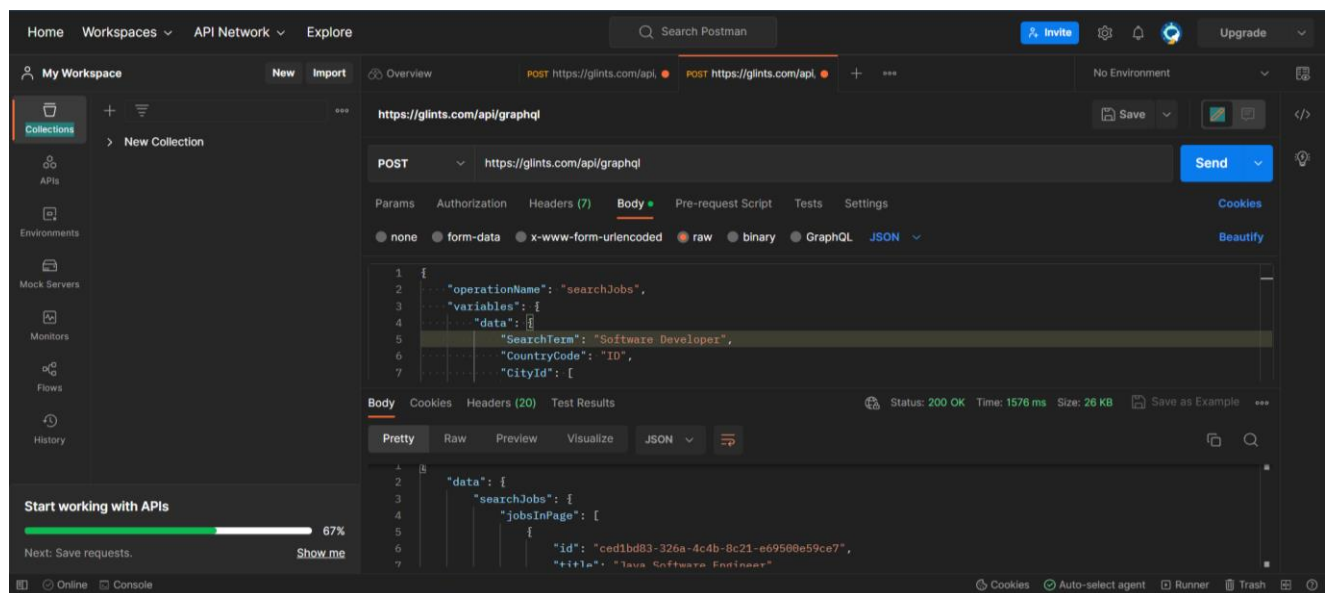


#### 4. Pengerjaan

- a) Menginput URL API ke dalam Postman dengan request method POST, kemudian pada bagian “Body” diisi dengan isi payload, untuk mendapatkan data yang diinginkan.



Dari Body tersebut, kita dapat mengubah variabel sesuai yang kita inginkan. Misalnya, dalam variabel SearchTerm, “Data Scientist” akan diubah menjadi “Software Developer” kemudian tekan “Send” maka hasil diperoleh seperti berikut:



- b) Melakukan Proses Request pada Visual Studio Code

Akan dijalankan Code Snippet (Python Request) pada website Postman pada Visual Studio Code seperti berikut:

```
import requests
import json

url = "https://glints.com/api/graphql"

payload = json.dumps({
  "operationName": "searchJobs",
  "variables": {
    "data": {
      "searchTerm": "Data Scientist",
      "countryCode": "ID",
      "cityId": {
        "id": "28904"
      },
      "limit": 30,
      "offset": 90,
      "includeExternalJobs": True,
      "source": {
        "NATIVE",
        "SUPER_POWERED"
      },
      "searchVariant": "QUERY_EXPANSION"
    }
  },
  "query": "query searchJobs(data: JobSearchConditionInput!) {\\n  searchJobs(data: $data) {\\n    jobsInPage {\\n      id\\n      title\\n      isRemote\\n      status\\n      createdAt\\n      updatedAt\\n      isActive\\n      isHiring\\n      isHot\\n      shouldShowSalary\\n      salaryEstimate\\n      company {\\n        id\\n        name\\n        logo\\n        __typename\\n      }\\n      citySubDivision {\\n        id\\n        name\\n        __typename\\n      }\\n      city {\\n        id\\n        name\\n        __typename\\n      }\\n      country {\\n        code\\n        name\\n        __typename\\n      }\\n      category {\\n        id\\n        name\\n        __typename\\n      }\\n      salaries {\\n        minYearsOfExperience\\n        maxYearsOfExperience\\n        source\\n        __typename\\n      }\\n    }\\n  }\\n}"
})

headers = {
  "Content-Type": "application/json"
}

response = requests.request("POST", url, headers=headers, data=payload)
print(response.text)
```

```
{
  "data": {
    "searchJobs": {
      "jobsInPage": [
        {
          "id": "9a3c14dc-3ae0-4130-867e-13e188545fcd",
          "title": "Data Analyst",
          "isRemote": false,
          "status": "OPEN",
          "createdAt": "2023-03-16T02:00:22.176Z",
          "updatedAt": "2023-04-06T01:09:21.987Z",
          "isActive": true,
          "isHiring": false,
          "isHot": false,
          "shouldShowSalary": true,
          "salaryEstimate": null,
          "company": {
            "id": "dea7f014-6b44-450b-acbd-a53562209c25",
            "name": "Eduwork.id",
            "logo": "a3a0b0a3454e5dab98d591e9361dd6e0.jpg",
            "__typename": "Company"
          },
          "citySubDivision": {
            "id": "90",
            "name": "Sunter",
            "__typename": "CitySubDivision"
          },
          "city": {
            "id": "28904",
            "name": "Jakarta",
            "__typename": "City"
          },
          "country": {
            "code": "ID",
            "name": "Indonesia",
            "__typename": "Country"
          },
          "category": {
            "id": "6",
            "name": "Product Management",
            "__typename": "JobCategory"
          },
          "salaries": [
            {
              "minYearsOfExperience": 3,
              "maxYearsOfExperience": 5,
              "source": "NATIVE",
              "__typename": "JobSalary"
            }
          ],
          "id": "743bb9d-bef9-4ebe-9451-afe672779fc2",
          "title": "Data Engineer",
          "isRemote": false,
          "status": "OPEN",
          "createdAt": "2023-02-27T07:22:03.756Z",
          "updatedAt": "2023-04-06T01:09:21.987Z",
          "isActive": true,
          "isHiring": false,
          "isHot": false,
          "shouldShowSalary": true,
          "salaryEstimate": null,
          "company": {
            "id": "e0e388f2-c503-460c-ab0c-a15b58f74d2c",
            "name": "PT Data Bina Solusindo",
            "logo": "abe0124457bf213c3b6096a1bf9c927.jpg",
            "__typename": "Company"
          },
          "citySubDivision": {
            "id": "90",
            "name": "Sunter",
            "__typename": "CitySubDivision"
          },
          "city": {
            "id": "28904",
            "name": "Jakarta",
            "__typename": "City"
          },
          "country": {
            "code": "ID",
            "name": "Indonesia",
            "__typename": "Country"
          },
          "category": {
            "id": "6",
            "name": "Product Management",
            "__typename": "JobCategory"
          },
          "salaries": [
            {
              "minYearsOfExperience": 3,
              "maxYearsOfExperience": 5,
              "source": "NATIVE",
              "__typename": "JobSalary"
            }
          ]
        }
      ]
    }
  }
}
```

c) Mengganti tipe data ke dalam bentuk JSON

```
# Ganti tipe ke JSON
import json

json_data = json.loads(response.text)
json_data
```

```
{
  "data": {
    "searchJobs": {
      "jobsInPage": [
        {
          "id": "9a3c14dc-3ae0-4130-867e-13e188545fcd",
          "title": "Data Analyst",
          "isRemote": false,
          "status": "OPEN",
          "createdAt": "2023-03-16T02:00:22.176Z",
          "updatedAt": "2023-04-06T01:09:21.987Z",
          "isActive": true,
          "isHiring": false,
          "isHot": false,
          "shouldShowSalary": true,
          "salaryEstimate": null,
          "company": {
            "id": "dea7f014-6b44-450b-acbd-a53562209c25",
            "name": "Eduwork.id",
            "logo": "a3a0b0a3454e5dab98d591e9361dd6e0.jpg",
            "__typename": "Company"
          },
          "citySubDivision": {
            "id": "90",
            "name": "Sunter",
            "__typename": "CitySubDivision"
          },
          "city": {
            "id": "28904",
            "name": "Jakarta",
            "__typename": "City"
          },
          "country": {
            "code": "ID",
            "name": "Indonesia",
            "__typename": "Country"
          },
          "category": {
            "id": "6",
            "name": "Product Management",
            "__typename": "JobCategory"
          },
          "salaries": [
            {
              "minYearsOfExperience": 3,
              "maxYearsOfExperience": 5,
              "source": "NATIVE",
              "__typename": "JobSalary"
            }
          ],
          "id": "743bb9d-bef9-4ebe-9451-afe672779fc2",
          "title": "Data Engineer",
          "isRemote": false,
          "status": "OPEN",
          "createdAt": "2023-02-27T07:22:03.756Z",
          "updatedAt": "2023-04-06T01:09:21.987Z",
          "isActive": true,
          "isHiring": false,
          "isHot": false,
          "shouldShowSalary": true,
          "salaryEstimate": null,
          "company": {
            "id": "e0e388f2-c503-460c-ab0c-a15b58f74d2c",
            "name": "PT Data Bina Solusindo",
            "logo": "abe0124457bf213c3b6096a1bf9c927.jpg",
            "__typename": "Company"
          },
          "citySubDivision": {
            "id": "90",
            "name": "Sunter",
            "__typename": "CitySubDivision"
          },
          "city": {
            "id": "28904",
            "name": "Jakarta",
            "__typename": "City"
          },
          "country": {
            "code": "ID",
            "name": "Indonesia",
            "__typename": "Country"
          },
          "category": {
            "id": "6",
            "name": "Product Management",
            "__typename": "JobCategory"
          },
          "salaries": [
            {
              "minYearsOfExperience": 3,
              "maxYearsOfExperience": 5,
              "source": "NATIVE",
              "__typename": "JobSalary"
            }
          ]
        }
      ]
    }
  }
}
```

d) Melakukan pembentukan data frame dengan module Pandas

## Python

5 rows x 36 columns