

Escuela de Ingeniería de Datos



Trabajo Final

⇒ ROCKINGDATA



Trabajo Final

Descargar dataset:
sadas



Consigna:

Deben aplicar al dataset compartido los temas vistos en el curso de Data Engineering y subirlos a su repo para que lo pueda ver el “honorable” equipo de evaluación de trabajos finales

Por ejemplo

limpieza de datos

Normalizacion

Conexion a datos.

Transformaciones



Deben entregar:

documento de jupyter

Codigo SQL

Documento que explique que hicieron y cómo.



Trabajo Final



Fecha de Entrega: Domingo 15 de agosto de 2021 a las 23:59hs
Enviar aviso a Cinthia de que ya se completó el trabajo final y el link al repo correspondiente.

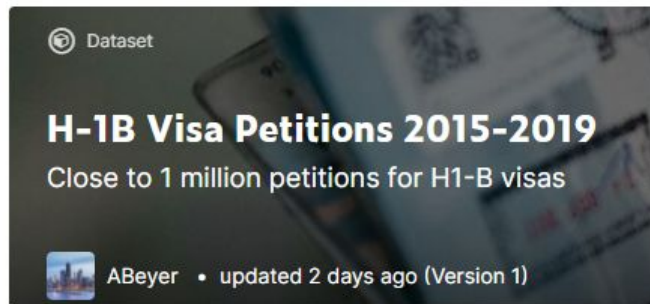


Certificado: Se entregará como constancia de haber completado el entrenamiento
- Requisito 100% de asistencia



¿Como es el Dataset?

<https://www.kaggle.com/abrambeyer/h1b-visa-petitions-20152019>



CASE_NUMBER,CASE_STATUS,EMPLOYER_NAME,SOC_NAME,SOC_CODE,JOB_TITLE,FULL_TIME_POSITION,PREVAILING_WAGE,WORKSITE_CITY,WORKSITE_STATE_ABB,YEAR,WORKSITE_STATE_FULL,WORKSITE
I-200-09285-514415,CERTIFIED,IT CATS LLC,COMPUTER SYSTEMS ANALYSTS,15-1121,SENIOR BUSINESS
ANALYST,Y,109116.8,"NORTH WALES","PA,2015,PENNSYLVANIA,"NORTH WALES, PENNSYLVANIA"
I-200-09341-564186,CERTIFIED-WITHDRAWN,"XPERT TECHNOLOGIES, INC.,""COMPUTER OCCUPATIONS, ALL
OTHER*","15-1799,QUALITY ASSURANCE ANALYST,Y,58282,CAMP HILL,PA,2015,PENNSYLVANIA,"CAMP HILL, PENNSYLVANIA"
I-200-09342-274341,CERTIFIED,"ISR INFO WAY, INC.,""SOFTWARE DEVELOPERS, APPLICATIONS",15-1132,SR. SOFTWARE
DEVELOPER,Y,84906,CHICAGO,IL,2015,ILLINOIS,"CHICAGO, ILLINOIS"



Links de referencias y extras

Hacé clic en cada enlace y te llevará al sitio web.

<https://vaex.io/docs/>

<https://spark.apache.org/docs/latest/api/python/>

<https://www.linux.org/>

<https://phoenixnap.com/kb/linux-commands-cheat-sheet>

https://es.wikipedia.org/wiki/Distribuci%C3%B3n_Linux

<https://www.techrepublic.com/article/microsoft-may-be-the-worlds-largest-open-source-contributor-but-developers-dont-yet-care/>

¡Muchas Gracias!

