

Udemy Data science Courses Dataset

1. Udemy เป็นแพลตฟอร์มคอร์สเรียนออนไลน์ ในหลากหลายหมวด รวมไปถึง Data Science ซึ่งทาง Data ที่ได้ถูกรวบรวม โดย Yasir Raza และมีการอัปเดตครั้งล่าสุดในปี 2022

1. ก่อนอื่นนะ เรามีการสำรวจ Data ที่ได้มา และทำความสะอาด Data โดยเราพบว่า ในส่วนของ Course Level โดยปกติจะมี All Levels , Beginner , Intermediate , Expert แต่มีทั้งหมด 6 คอร์สด้วยกัน ที่มีลักษณะแตกต่างจากเพื่อนเราเลยทำการครีเอต

1. ต่อมานะ เรามีการสำรวจ แต่ละคอลัม ว่าเรื่องไหนน่าสนใจ และสามารถนำมาเชื่อมโยงกันได้ จึงได้มาดังนี้ละ

1. โดยทางกลุ่มเรา นะครับ มีการนำ Rating ของอาจารย์แต่ละคน มาหาค่าเฉลี่ย เพื่อหาว่าชื่อของผู้สอน มีผลต่อคุณภาพของการสอนหรือไม่ ซึ่งใน bar chart นี่นะ จะมีการเรียงลำดับผู้สอนจากจำนวนคอร์สที่อาจารย์สอนจากน้อยไปมากละ

1. นอกเหนือจากหาค่าเฉลี่ยแล้ว ทางกลุ่มเราได้มีการนำตัวค่าเฉลี่ย มาเทียบกับ max ของอาจารย์ท่านนั้นๆ จะได้ดัง bar chart นี่นะ ละ สังเกตดูมัยละว่าจะมีอาจารย์บางท่าน ที่มีค่าเฉลี่ยต่ำ หรือแตกต่างจากค่า max จนเกินไป เช่น Spotle Learn เราเลยทำการเอาแค่คอร์สส่วนของ Spotle Learn พบว่าสาเหตุที่ทำให้ค่าเฉลี่ยต่ำมากๆ เนื่องจากมีคอร์สจำนวนมาก ที่มีค่า Rating เท่ากับ 0.0

1. ใช่ว่า เราเลยทำการตรวจสอบความสัมพันธ์ของ Rating กับ Review พบว่าคอร์สไหนที่มีจำนวนรีวิวเท่ากับ 0 จะมี Rating เท่ากับ 0 เช่นกัน พวกเราเลยอนุมานว่าคอร์สไหนที่มีรีวิวเท่ากับ 0 ถือว่ายังไม่มีคนที่เข้ามาเรียนคอร์สนี้ และ ยังไม่นับมาหาค่าเฉลี่ยด้วย

1. และนี่คือ bar chart หลังจากทำการจบคอร์สที่มีรีวิวเป็น 0 ทิ้งแล้วจะสังเกตว่าตัว bar chart มีความ stable มากขึ้น

1.อีกทั้งเรายังทำแบบสอบถามเพื่อสนับสนุนว่าที่เราคิดนั้นถูกต้องรึป่าว จากแบบสำรวจพบว่าส่วนใหญ่ซึ่งอาจารย์ยังมีผลต่อการเลือกคอร์สเรียนแต่ที่มีผลมากที่สุดคือ ราคา และ รีวิวจากผู้ที่เคยเรียนแล้ว

1.นอกจากนี้กลุ่มเรายังมีการสำรวจ Programming Languages ที่สอนได้ออกมาดังนี้ โดย 6 อันดับแรกที่เราเห็นได้ชัดได้แก่ Python,C,R,Golang,Java,SQL

1.พวกเราเลยมีข้อสงสัยว่าทำไม 6 ภาษานี้จึงมีเปอร์เซนต์สูง และเมื่อเทียบจากสถิติของ Programming Language ที่นิยมใช้มากที่สุด เลยพบว่าสิ่งนี้น่าจะเป็นสาเหตุที่ทำให้มีเปอร์เซนต์สูง อีกทั้งยังคาดการณ์ว่าภาษาไหนจะมีการนิยมใช้ในปี 2024 พบว่าตัวPythonนิยมใช้อย่างต่อเนื่อง

1.กลุ่มพวกเราเลยมาหารว่าเพราะอะไร Python จึงถูกนิยมใช้พบว่าตัวPython มีsyntaxที่เข้าใจและศึกษาได้ง่าย สำหรับผู้เขียนโปรแกรมทุกระดับ