



Labs PI03: Analytics

¿QUE PASO?:

Los MOOCs cursos masivos abiertos y online, por sus siglas en inglés han revolucionado el mundo de la educación desde principios de la década pasada, cuando el profesor Sebastian Thrun comenzó con la transmisión online de su curso introductorio a la Inteligencia Artificial. Poco tiempo después, Thrun fundó Udacity.

Con el pasar de los años se han ido sumando otras plataformas como edX y Coursera. Brindando servicios similares: acceso a contenido específico, de calidad y de manera práctica, desde la comodidad del hogar. Muchas de estas plataformas tienen contenido gratuito mientras que el modelo de negocio en general se basa ya sea en el pago de suscripciones recurrentes para acceso general o únicas, para acceder a certificaciones o a cursos premium.

Con el aumento de la popularidad de los MOOCs, no solo han aparecido nuevas plataformas privadas como las mencionadas anteriormente, sino que también muchas universidades y organizaciones sin fines de lucro han sumado a la oferta haciendo el mercado mucho más competitivo.

En este contexto, resulta imperante para cada plataforma, ajustar sus modelos de negocio, los cursos y el contenido que se ofrece en los mismos para lograr captar y retener a la mayor cantidad de clientes.

Coursera Inc.

es un proveedor masivo de cursos abiertos en línea con sede en Estados Unidos fundado en 2012 por los profesores de ciencias de la computación de la Universidad de Stanford Andrew Ng y Daphne Koller. Coursera trabaja con universidades y otras organizaciones para ofrecer cursos, certificaciones y títulos en línea en una variedad de temas.

Precio de las acciones: COUR (NYSE) \$15.51 +0.02 (+0.13%) - 20 de mayo, 10:12 AM EDT

Fundado: 2012

Sede: California, Estados Unidos

Usuarios: 92 millones (2021)

Empleados: 1,138 (diciembre 2021)

Disponible en: Multilingüe (40)

Fundadores: Andrew Ng, Daphne Koller

Con este análisis es posible evaluar qué cursos / temáticas son más rentables, lo cual contempla no solo cantidad, sino también precio.

Edx

es una plataforma de cursos abiertos masivos en línea, conocidos por sus siglas en inglés como MOOC (Massive Online Open Course), basada en software de código abierto. Fue fundada en el Instituto Tecnológico de Massachusetts (MIT) y la Universidad de Harvard en mayo de 2012 para hospedar cursos en línea de nivel universitario de un amplio rango de disciplinas, para todo el mundo sin costos para propiciar la investigación y el aprendizaje. Cada una de estas dos instituciones contribuyen con 30 millones de dólares de recursos para este proyecto sin ánimo de lucro. El curso prototipo con el cual empezó el programa fue Circuitos y Electrónica, que comenzó en diciembre de 2011, a través de MITx, el programa en línea y masivo del MIT.

EdX tiene más de 10 millones de usuarios registrados y es considerado el segundo proveedor más grande de MOOC en el mundo, después de Coursera que cuenta con 23 millones de usuarios registrados.

A finales de 2016, edX ofrecía 1300 cursos de los cuales 500 fueron lanzados en 2016. Del total de cursos que ofrece, 53 son pagados, conocidos como cursos de educación profesional, y 13 son elegibles para créditos universitarios. De acuerdo a una investigación realizada por Class Central, el costo promedio de un certificado de edX es de USD\$53.

Actualmente hay más de 130 colaboradores entre institutos, organizaciones sin ánimo de lucro, corporaciones y organizaciones internacionales que ofrecen o planean ofrecer más de 500 cursos en el sitio web de edX.

Tipo: Educación en línea

Sin fines de lucro

Registro Obligatorio

Idiomas disponibles Inglés, español, francés, Hindi y mandarín

Estado actual Activo

Gestión

Propietario MIT y Harvard University

Lanzamiento Mayo de 2012

Estadísticas

Usuarios registrados 35 millones (febrero de 2021)¹

Ranking Alexa Decrecimiento 1,178 (junio de 2017)²

udemy

Udemy.com es una plataforma de aprendizaje en línea. Está dirigido a adultos profesionales. A diferencia de los programas académicos MOOC conducidos por tradicionales cursos de trabajo creados por las universidades, Udemy utiliza contenido de creadores en línea para vender y así conseguir ganancias. Udemy proporciona herramientas para que los usuarios puedan crear un curso, promocionarlo y ganar dinero con gastos de inscripción de los estudiantes.

Ninguno de los cursos de Udemy son actualmente equivalentes a la obtención de un título universitario. Los estudiantes toman cursos en gran parte simplemente como medio para mejorar sus habilidades. Algunos cursos generan crédito hacia certificación técnica. Udemy ha hecho un esfuerzo especial para atraer a empresas que buscan crear cursos de trabajo especialmente para empleados de su compañía.

En 2018, había más de 65.000 cursos disponibles en el sitio web.

Tipo; Sitio web

Industria Educación

Forma legal: empresa de capital abierto

Fundación 2010

Sede central San Francisco (Estados Unidos)

Sitio web www.udemy.com

Nuestra PM se dirigió a nosotros con un nuevo ticket de trabajo

Una **startup de tecnología**, está interesada en sumarse al mercado de cursos online, pero de una manera eficiente, por lo que compró datasets de posibles competidores para analizar y sacar conclusiones de los datos recolectados.

Ellos solicitan segmentar el nivel de ventas según precio, idioma, nivel y rating, de cada curso para analizar qué tanto influyen dichas variables en la demanda del producto vendido.

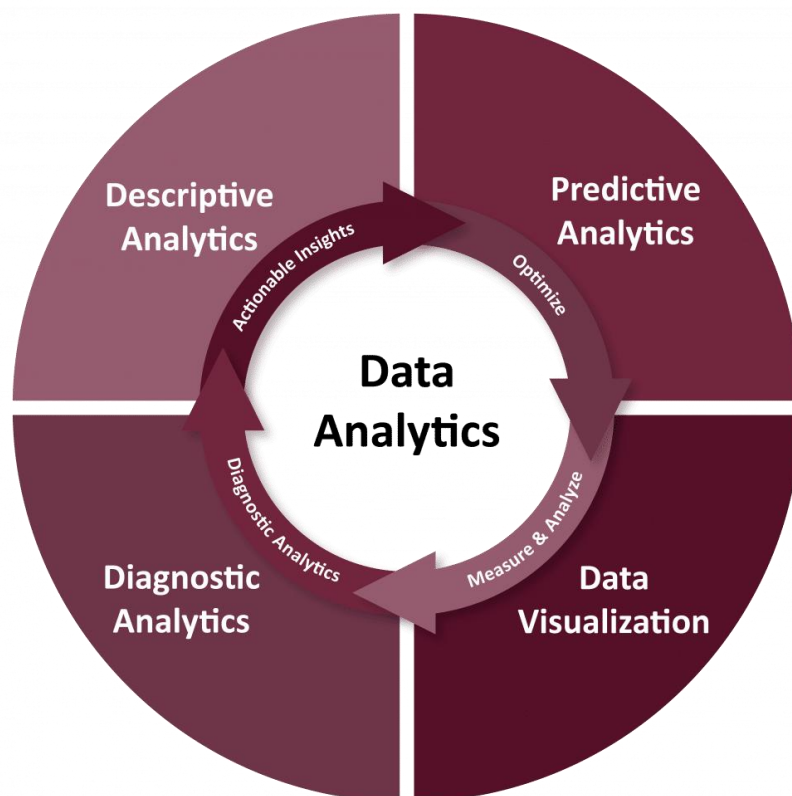
Por otra parte se nos solicita un **WordCloud de las palabras clave que más se repiten dentro del título**. (Se puede añadir otras variables de nuestro interés).

Por último, se nos pide **una demo** en un rango de tiempo de **no más de 10 min** donde presentamos las funcionalidades del **dashboard** y las **conclusiones / recomendaciones** de nuestra parte.

Con el fin de monitorear la eficacia de los objetivos de la empresa, se le pide **establecer _al menos_ 1 KPI** producto de su análisis y que el mismo se pueda **visualizar** en un **dashboard**.

Por consiguiente en el presente readme, estamos describiendo cada uno de los pasos del proyecto PI03-DTS06 de Henry Students, demostrando una vez más lo aprendido durante el bootcamp en este caso bajo el rol de un Data Analyst.

Metodología a Seguir



Información a partir de Datos

EDA

El análisis exploratorio de datos se refiere al proceso crítico de realizar investigaciones iniciales sobre datos para descubrir patrones, detectar anomalías, probar hipótesis y verificar suposiciones con la ayuda de estadísticas resumidas y representaciones gráficas.

Exploración de conjuntos de datos

Resumen de datos

El conjunto de datos utilizados, fue dotado por el repositorio de la consigna, Aparte como plus use dos datasets adicionales descargados del sitio web de Coursera. Este conjunto de datos contiene principalmente 6 columnas y 890 datos del curso. La descripción detallada:

course_title : Contiene el título del curso.

course_organization : Indica qué organización está llevando a cabo los cursos.

courseCertificateType : Tiene detalles sobre cuáles son las diferentes certificaciones disponibles en los cursos.

course_rating : Tiene las calificaciones asociadas a cada curso.

course_difficulty : Habla sobre qué tan difícil o cuál es el nivel del curso.

coursestudentsenrolled : Tiene el número de alumnos que están matriculados en el curso.

Análisis Exploratorio de los datos (_Exploratory Data Analysis = EDA_):

Archivo notebook (.ipynb) de este repositorio

[reporte.ipynb](#)

el cual esta adecuadamente comentado.

Comenzamos con Tech skills

Las habilidades técnicas son las habilidades y conocimientos necesarios para realizar tareas específicas. Son prácticos y a menudo se relacionan de tecnología de la información, matemáticas o científicas. En este caso se muestra el conocimiento en lenguaje de **programación** (Python), Recolección (datasets) **Transformación** (EDA, Carga de datos y revisión básica, Aplicando el EDA comprobamos los datos antes de hacer cualquier suposición. lo que me Permite identificar errores obvios, así como comprender mejor los patrones en los datos, gracias a ello trabajare este analisis con el Dataframe numero df_2, df_3 y df_4 como tambien con datos de la web y no incluye el primer Dataframe por la razon obvia de no tener los datos que requiero. Procedere a revisar valores atípicos o sucesos anómalos.) y **Visualización** (Dashboard, Metricas, KPIs), etc.

Análisis y Conclusiones del Dashboard:

Comenzamos esta tarea realizando el EDA con la consiguiente limpieza y transformación de datos, ingestado 4 archivos .csv en principio. Y dos .csv adicionales como plus donde en primer lugar para enriquecer mi comprensión y criterio indague acerca de los MOOCs y sobre cada una de las empresas provistas en los diferentes datasets, donde de una manera aplastante se ubica en primer lugar COURSERA con varias características que la sitúan de manera firme como la tendencia mas fuerte a considerar en este mercado, así fui repasando cada uno de los datos y estableciendo metricas de las cuales resaltan 1.454.711 inscritos o matriculados (Uno de los KPIs de mi analisis), con este proveedor masivo, este gigante fundado en 2012 por los profesores de ciencias de la computación de la Universidad de Stanford Andrew Ng y Daphne Koller. trabaja con universidades y otras

organizaciones para ofrecer cursos, certificaciones y títulos en línea en una variedad de temas. pero lo que realmente destaca es que como proveedor de cursos maneja en su folder las mas prestigiosas y mejores instituciones en este campo (otro KPIs clave), en este contexto los datasets analizados no solo aportaron metricas de valor sino que tambien nos situaron en un contexto temporal sin precedentes, lo cual supone el KPIs de mayor importancia porque nos comunica inmediatamente no solo en lo ¿QUE PASO? si no que nos transmite ¿Que esta pasando? de una manera clara.

Por lo tanto manejando ahora no solo los KPIs si no las diversas metricas, que no dejan de ser importante entre nuestro analisis nos prepara claramente para dar respuesta a nuestra mayor responsabilidad en el contexto ¿Que esta pasando? y ¿Por que?.

Evidentemente el conjunto de datos analizados son del año 2020 (lo que nos situa de manera temporal en el inicio de la pandemia) por tal motivo de una manera responsable, entendemos que tal comportamiento en el incremento exponencial de cada una de las metricas en este campo. Obedece directamente al comportamiento causal del contexto. por lo tanto es hora de pensar como mantener este singular crecimiento mientras se sumergue la situación cada día mas en una normalidad, que supone de manera abrupta un cambio en esta tendencia de crecimiento.

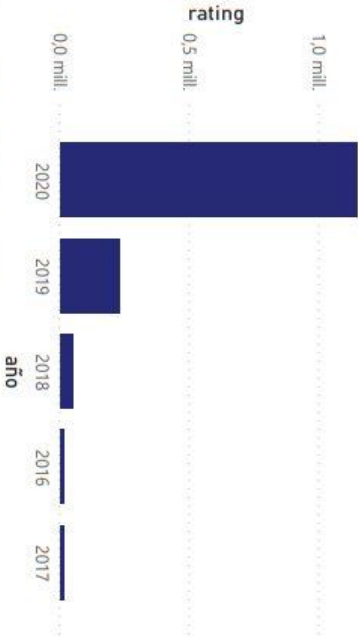
Cuando aplicamos la media como tendencia estadística a cada una de nuestras metricas nos damos por enterados en este acercamiento mas fiel a las tendencias normales de comportamiento del mercado habitual, que la meta en consecuencia es no solo mantenerse en los valores de la media estadística, si no aplicar distintas acciones en cada uno de los niveles estrategicos de Nuestra PM, lo cual garantiza una manera eficiente de entonizarse en este mercado, el margen de error depende directamente de la fiabilidad de la compra de los datasets de posibles competidores pues los analisis y las conclusiones dependen en gran medida de los datos recolectados.

Se segmentaron las metricas para el nivel de ventas según **precio, idioma, nivel y rating** de cada curso, el analisis directamente muestra que tanto influyen dichas variables en la demanda del producto vendido. lo cual ya ha sido expuesto previamente.

Conclusion: Seria de bastante ayuda contar con datos mas actualizados en los datasets para hacer una proyeccion mas realista y cotejar los resultados de este analisis hasta el 2020 (situandolo como un analisis Descriptivo), con este lapso de dos años lo que evindentemente certificara la fiabilidad en este caso del analisis en cuestion, y nos permitira una proyeccion para conocer ¿Que esta pasando? y ¿Por que? de una forma mas actual y ver

MOOCs

rating por año y últimos 5 años



Suma de cantidad de cursos



Column 1	Suma de cantidad de cursos	nivel
	621	Introductory
	267	Intermediate
	87	Advanced
Total	975	

KPI por ¿Que paso?

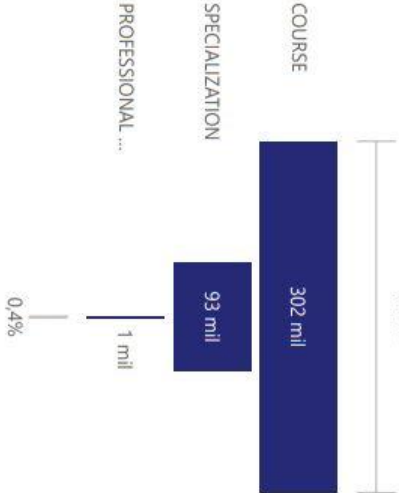
1454711

\$49.00 / 270 COURSES

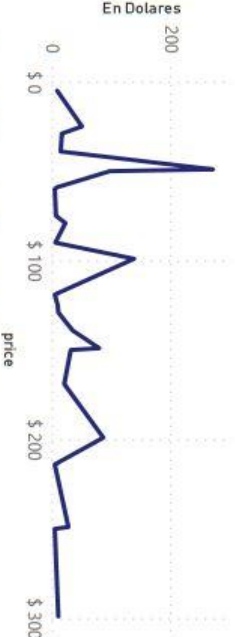
valor Atipico y ultimo año



Tipos por course_Certificate_type



En Dolares por price



orden por Mejor Proveedor de cursos

