

Abrir en la aplicación ↗



Buscar



✦ Historia exclusiva para miembros

La cartera que me consiguió un trabajo de científico de datos

Alerta de spoiler: fue sorprendentemente fácil (y gratuito) hacer



Matt Chapman Seguir

Publicado en Hacia la ciencia de datos

Lectura de 10 minutos 24 de marzo de 2023



Escuchar



Compartir



Más



Foto de [John Schnobrich](#) en [Unsplash](#)

Conseguir un trabajo de científico de datos es difícil.

Ya no estamos en 2015: no basta con conocer algunas funciones de pandas y poner las palabras "Big Data" en su resumen. La competencia por los puestos más altos es feroz. En un reciente

Mientras revisaba la bolsa de empleo de LinkedIn, me costó encontrar un puesto de científico de datos con sede en Londres con menos de 100 solicitantes.

La buena noticia es que esta competencia no se debe a la falta de empleo. Incluso en 2023, la ciencia de datos seguirá siendo un campo de rápido crecimiento, y la Oficina de Estadísticas Laborales de EE. UU. ha estimado que el número de empleos de científicos de datos crecerá un 36 % entre 2021 y 2031 [1].

El desafío, entonces, no es que no haya trabajos en ciencia de datos: ¡hay toneladas!

Más bien, es que una gran cantidad de personas están tratando de ingresar a la industria, lo que hace cada vez más difícil destacar entre la multitud y conseguir ese lucrativo primer puesto.

La solución: crear un portafolio de ciencia de datos En el mercado

laboral actual, estoy convencido de que una de las mejores formas de diferenciarse de la competencia es mediante la creación de un portafolio personal de proyectos personales de ciencia de datos para mostrar sus habilidades y experiencia. Esto es especialmente importante si no tiene mucha experiencia comercial trabajando en roles de ciencia de datos. Para tomar prestadas las palabras del científico de datos Will Stanton [2]:

Si no tienes experiencia como científico de datos, entonces es absolutamente necesario que realices proyectos independientes.

En este artículo, te guiaré a través del proceso que seguí para crear mi portafolio de proyectos personales e intentaré darte un poco de inspiración sobre cómo puedes crear uno para ti.

Si la idea de crear un portafolio le parece un poco desalentadora, no se preocupe: es sorprendentemente fácil (y gratuito) poner algo en funcionamiento, y no necesita ningún conocimiento de desarrollo web para comenzar.

Mi experiencia (también conocido como por qué creé mi portafolio)

Mi viaje hacia la ciencia de datos comenzó hace relativamente poco tiempo. Mis primeros trabajos recién salidos de la escuela/universidad fueron en campos como gestión de proyectos, ventas y marketing. Cuando decidí hacer el cambio a la ciencia de datos en 2020, tenía muy poca experiencia práctica trabajando con datos y no tenía una licenciatura técnica en un campo como Ciencias de la Computación o Estadística.

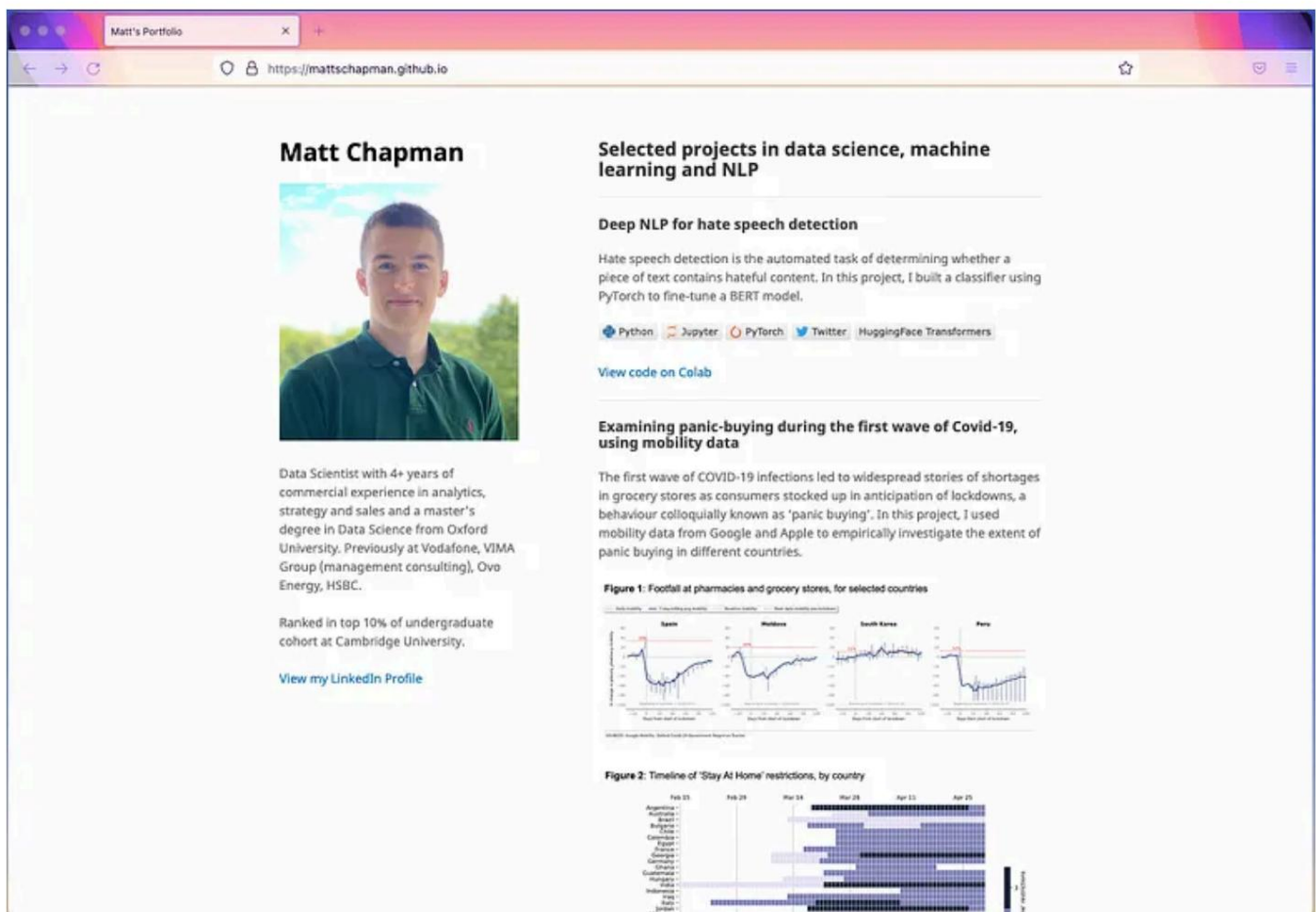
Vaya. Ese no es un buen punto de partida.

En mi publicación anterior, compartí cómo abordé esto tomando varios cursos en línea, estudiando una maestría en ciencia de datos y realizando un par de pasantías en ciencia de datos. A lo largo del camino, trabajé en muchos proyectos diferentes y construí un amplio repertorio de habilidades relacionadas con los datos.

Sin embargo, para los primeros trabajos a los que postulé, descubrí que me costaba poner un pie en la puerta: sentía como si los reclutadores ni siquiera estuvieran revisando mi solicitud, a pesar de que cumplía con todos los requisitos mínimos. En ese momento, esto fue increíblemente frustrante. Creí que tenía mucho que ofrecer, pero parecía que las conversaciones se estaban cerrando incluso antes de comenzar.

Para empeorar las cosas, aprendí que los reclutadores suelen dedicar tan solo entre 6 y 8 segundos a revisar un CV para una oferta de trabajo con exceso de suscripciones [3]. Eso me convenció de que necesitaba una forma más rápida de comunicar mis habilidades y destacarme, especialmente porque no tenía mucha experiencia como científico de datos.

Para abordar este problema, decidí crear un portafolio en línea como una forma de mostrar mi experiencia en un formato visualmente atractivo y rápidamente digerible.



Mi cartera de ciencia de datos: <https://mattchapman.github.io/>. Fuente de la imagen: propia del autor.

Luego, una vez que creé mi portafolio, incluí un enlace en la parte superior de mi currículum, para que sea muy fácil de encontrar para los reclutadores:



El encabezado de mi currículum/CV. Me gustó la idea de tener un enlace o código QR en el que se puede hacer clic que llevaría a los empleadores directamente a mi cartera y me ayudaría a destacar entre la multitud. Fuente de la imagen: propia del autor.

Cómo creé mi portafolio

Al crear mi portafolio, seguí algunos principios básicos:

1. Incluya proyectos que tengan algo genuinamente único e interesante.
a ellos
2. Sea breve y sencillo
3. Hazlo bonito
4. No pierda el tiempo en el lado del "desarrollo web"

En el resto de este artículo, analizaré brevemente cada uno de ellos y explicaré cómo me ayudaron a crear una cartera ganadora.

Incluya proyectos que tengan algo verdaderamente único e interesante.

Esto puede parecer obvio, pero es un punto tan importante que vale la pena repetirlo:

No incluyas contenido que todos los demás incluyen.

Si el objetivo de su cartera es diferenciarse, ¿por qué se molestaría en incluir proyectos genéricos que otros ya han hecho hasta el cansancio? Este es un punto del que se hizo eco Jeremie Harris en Las 4 formas más rápidas de no ser contratado como científico de datos [4], donde dijo: _____

Es difícil pensar en una forma más rápida de incluir su currículum en la pila de "definitivamente no" que presentar el trabajo que realizó en conjuntos de datos triviales de prueba de concepto entre sus proyectos personales destacados.

En caso de duda, aquí tienes algunos proyectos que te perjudican más de lo que te ayudan:

* Clasificación de supervivencia en el conjunto de datos del Titanic.

* Clasificación de dígitos escrita a mano en el conjunto de datos MNIST. _____

* Clasificación de especies de flores utilizando el conjunto de datos del iris.

Lo que esto significa es que si está realmente interesado en la visión por computadora, incluya un par de proyectos relacionados con eso. O, si simplemente no te cansas de los problemas de las series temporales, agrega algo sobre ese tema. Honestamente, no importa si está interesado en la optimización bayesiana o en el tejido de cestas: el punto es que asegúrese de que sea algo que realmente le parezca interesante y que crea que agregará valor a los demás. Personalmente, estoy muy interesado en problemas relacionados con el procesamiento del lenguaje natural y las ciencias sociales, así que intenté incluir proyectos relacionados con ellos. Pero, sinceramente, no importa tanto lo que incluyas, siempre y cuando puedas demostrar de forma convincente por qué añade valor.

Manténlo corto y simple

Nadie quiere perder tiempo en una página web larga o aburrida, y esto es especialmente cierto cuando se trata de portafolios en línea. Estás dando una instantánea, no escribiendo una tesis doctoral, y honestamente creo que perderás el tiempo si dedicas demasiado tiempo a escribir largas descripciones de proyectos y secciones de resultados.

Para apreciar este punto, tenemos que volver al propósito original de un portafolio, y en el fondo hay una idea simple:

El propósito de su cartera es poner un pie en la puerta, no conseguirle el trabajo.

En otras palabras, nunca recibirá una oferta de trabajo únicamente del fondo de su cartera. La realidad es que, incluso si a los reclutadores les gusta el aspecto de su cartera, igual tendrá que pasar por sus entrevistas y evaluaciones. El propósito del portafolio es simplemente brindarles una instantánea rápida de sus habilidades y mostrarles lo que puede hacer.

En mi portafolio, por ejemplo, mencioné siete proyectos en los que había trabajado anteriormente. Para seis de los siete proyectos, no escribí más de dos breves frases de descripción.

Sí, has leído bien: dos frases.

Aquí hay un ejemplo de una descripción de proyecto en mi cartera:

Deep NLP for hate speech detection

Hate speech detection is the automated task of determining whether a piece of text contains hateful content. In this project, I built a classifier using PyTorch to fine-tune a BERT model.



Python



Jupyter



PyTorch



Twitter

HuggingFace Transformers

[View code on Colab](#)

Un proyecto de ejemplo en mi cartera. Fuente de la imagen: propia del autor.

Como puede ver, no perdí el tiempo estableciendo el contexto y explicando los matices de los resultados. Claro, eso es muy importante. Pero si el posible empleador está interesado, le preguntará en la entrevista. En esta etapa, solo necesita una breve descripción que pueda despertar su interés y generar más discusiones más adelante. Si lo desea, puede incluso incluir un enlace a más descripciones en otro lugar; eso es lo que hice al incluir un enlace al código debajo de la descripción del proyecto. Pero en la página principal de su cartera, sólo necesita los detalles de alto nivel. Cualquier otra cosa es una pérdida de tiempo.

Hazlo bonito

Mi siguiente consejo es hacer que su portafolio sea visualmente atractivo.

Ponte en el lugar de un posible reclutador: has leído cientos (o incluso miles) de currículums insulsos y te cuesta recordar cómo luce el color verde porque ha pasado mucho tiempo desde que viste algo que no fuera negro, y -blanco. Luego, aparece un currículum/portafolio que, sosten la puerta de entrada, es bastante agradable de ver.

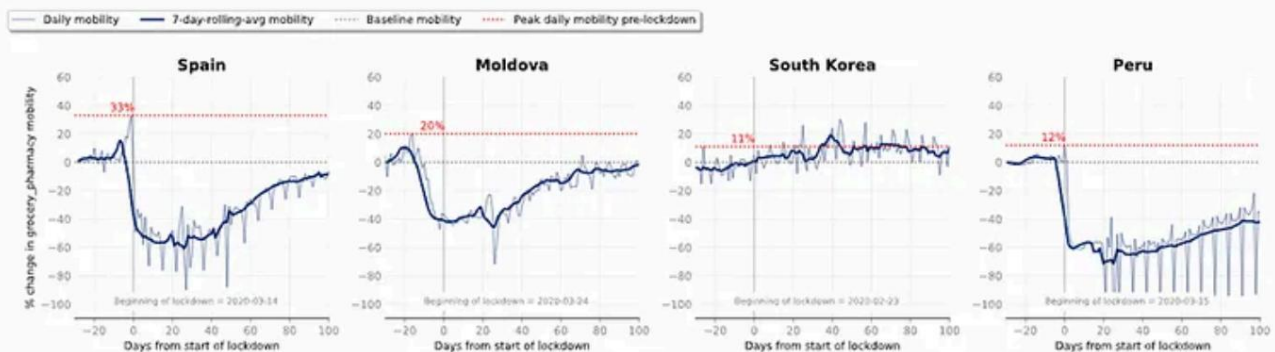
No es necesario ser un científico espacial (ni siquiera un científico de datos) para darse cuenta de que el reclutador disfrutará mirando su cartera.

Para que su portafolio sea bonito, mi principal consejo es incluir muchos gráficos/tramas interesantes. Por ejemplo, aquí está la descripción de otro proyecto que incluí en mi portafolio, que tiene algunos gráficos sobre las restricciones de Covid-19 y los movimientos de población durante los encierros:

Examining panic-buying during the first wave of Covid-19, using mobility data

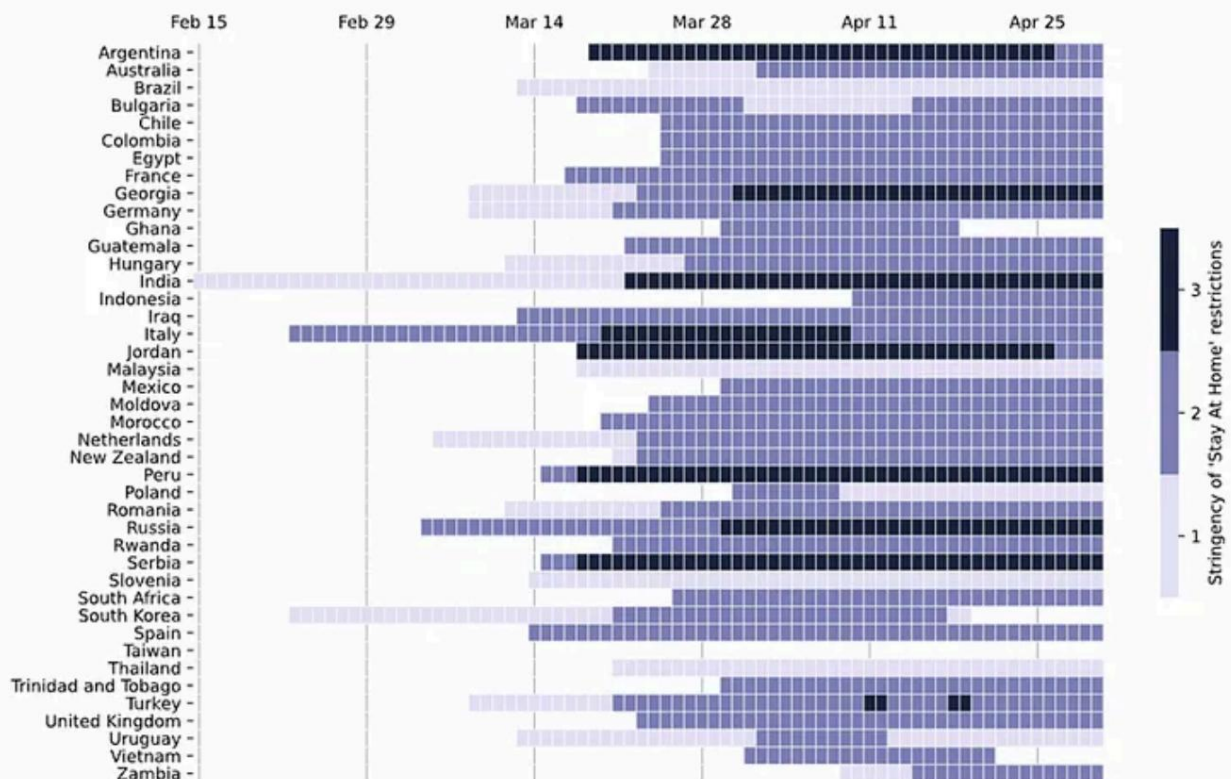
The first wave of COVID-19 infections led to widespread stories of shortages in grocery stores as consumers stocked up in anticipation of lockdowns, a behaviour colloquially known as 'panic buying'. In this project, I used mobility data from Google and Apple to empirically investigate the extent of panic buying in different countries.

Figure 1: Footfall at pharmacies and grocery stores, for selected countries



SOURCES: Google Mobility, Oxford Covid-19 Government Response Tracker

Figure 2: Timeline of 'Stay At Home' restrictions, by country



SOURCE: Oxford Covid-19 Government Response Tracker

View code on Colab

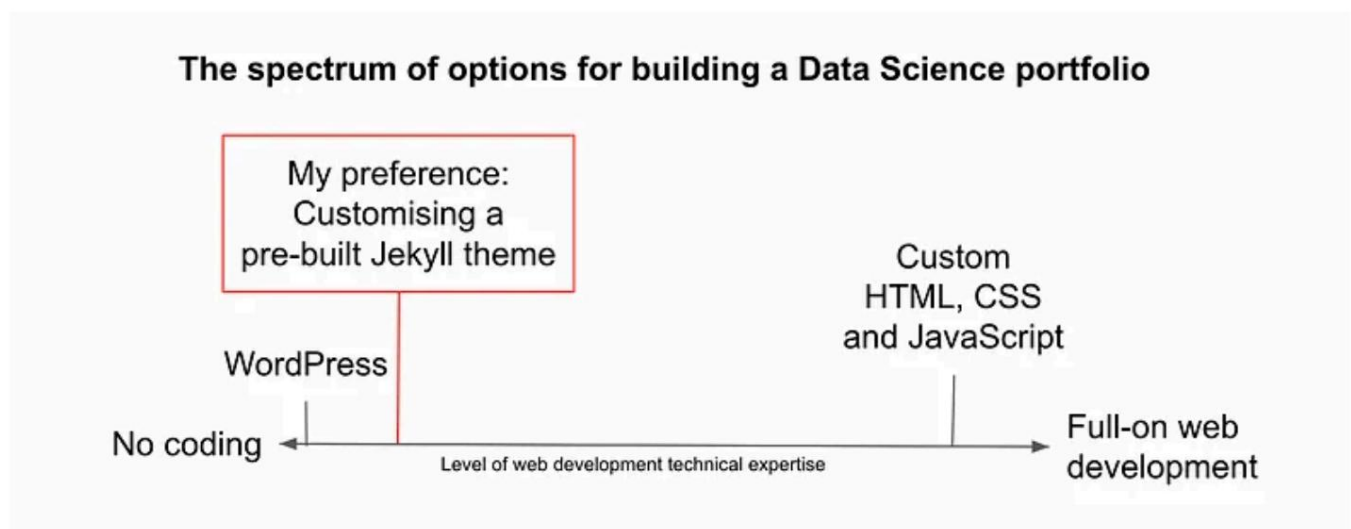
Una de las descripciones de proyectos en mi cartera. Fuente de la imagen: propia del autor.

Los gráficos muestran a los posibles empleadores que sé cómo comunicar ideas visualmente y crear imágenes bonitas. Esto puede parecer trivial, pero en realidad no lo es: como científico de datos, con frecuencia tendrás que comunicar tus hallazgos a partes interesadas no técnicas. Al incluir algunos gráficos en tu portafolio, estás mostrando a tus lectores que reconoces la importancia de la comunicación y que tienes las habilidades necesarias para ponerla en práctica. En otras palabras, es una manera fantástica de mostrar tus habilidades y diferenciarte en un campo abarrotado.

No pierdas tiempo innecesario en el desarrollo web

Mi último consejo es evitar dedicar demasiado tiempo a “construir” el sitio web que contendrá su cartera. Recuerde: está solicitando un trabajo como científico de datos, no como desarrollador web. No se le evaluará por sus habilidades en HTML y CSS; Se le evaluarán sus habilidades en ciencia de datos.

En consecuencia, a menos que tenga amplias habilidades previas en desarrollo web, le recomendaría optar por la opción de código bajo/sin código para crear su sitio web, por ejemplo utilizando una herramienta como WordPress o Webflow que le permita “arrastrar y soltar”. ”” secciones de su cartera y elija entre una variedad de plantillas prediseñadas.



El espectro de opciones para crear una cartera de ciencia de datos. Fuente de la imagen: propia del autor.

En mi caso, opté por algo un poco diferente y seguí esta excelente guía escrita por Ivanna Kacewica, quien utilizó el tema Minimal Jekyll para GitHub Pages para construir lo básico del sitio web [5]. Las principales razones por las que elegí esta opción.

donde (a) es gratis y (b) no requiere que incluya anuncios o la palabra "wordpress" en la URL.

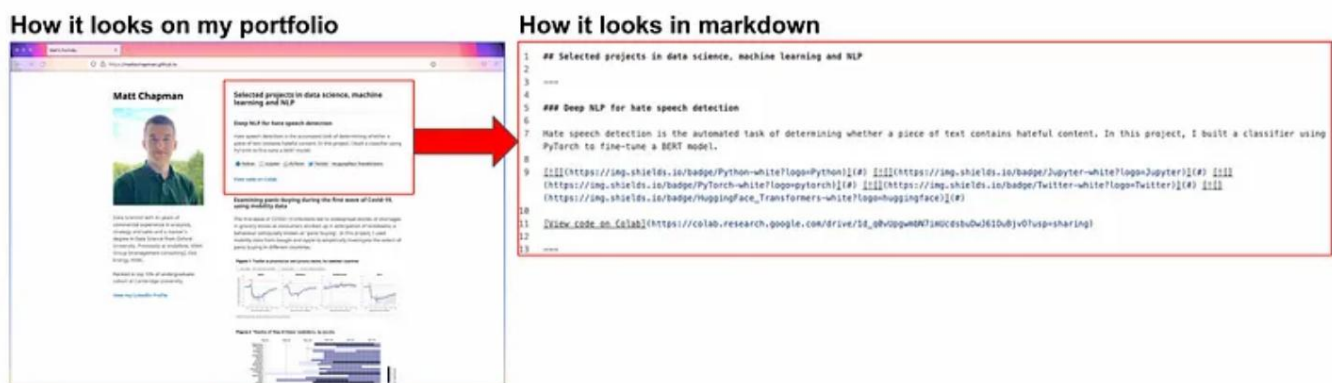
Una ventaja adicional del enfoque de Ivanna es que le permite alojar su cartera de forma gratuita en páginas de GitHub en la dirección nombre de usuario.github.io, donde nombre de usuario es su nombre de usuario (o nombre de organización) en GitHub. Dado que su nombre de usuario de GitHub está incluido en la URL, es muy fácil para las personas encontrar su cuenta de GitHub, lo cual es útil si tiene proyectos adicionales almacenados en GitHub.

No entraré en detalles sobre cómo usar el tema de Ivanna (ya que ella hace un excelente trabajo explicando las cosas), pero la idea básica de este enfoque es agregar texto e imágenes usando Jekyll y Markdown. Si no está familiarizado con Jekyll, es un marco simple para crear sitios web que no requiere codificación ni conocimientos de HTML/desarrollo web.

Con Jekyll, construyes la estructura básica del sitio web y luego, con Markdown, completas el sitio con tu texto e imágenes. Si no ha utilizado Markdown antes, no se preocupe:

Markdown es simplemente una forma realmente sencilla de formatear texto. Y cuando digo "simple", me refiero a realmente simple: para escribir un encabezado, por ejemplo, simplemente coloca un hashtag delante del texto que deseas que esté en tu encabezado.

Para hacer una viñeta, simplemente coloque un asterisco delante del texto que desea que esté en su lista. Es fácil.



Cómo agregar contenido usando Markdown. Fuente de la imagen: propia del autor.

Si desea ver todos los detalles de cómo personalicé el tema, consulte el archivo index.md en mi repositorio de GitHub.

Conclusión

Crear un portafolio en línea es una excelente manera de destacarse en el abarrotado mercado laboral de la ciencia de datos. En este artículo, he dado una idea de cómo creé mi portafolio personal y mis mejores consejos para crear el tuyo propio.

Déjame saber cómo te va y, si deseas recibir comentarios o consejos sobre tu portafolio, deja un enlace en los comentarios y me aseguraré de echarle un vistazo.

Si desea obtener más consejos sobre cómo generar ideas para un portafolio, puede consultar uno de mis otros artículos:

Cómo encontrar ideas únicas para proyectos de ciencia de datos que sean propias
Destaca la cartera

Olvidate de Titanic y MNIST: elige un proyecto único que desarrolle tus habilidades
y te ayude a destacar entre la multitud.

haciadatascience.com

Ah, una cosa más: comencé un

boletín informativo gratuito llamado AI in Five, donde comparto 5 viñetas cada semana sobre las últimas noticias sobre inteligencia artificial, consejos de codificación e historias

profesionales para científicos/analistas de datos. No hay exageraciones, no hay basura de "los datos son el nuevo petróleo" ni tweets de Elon: solo consejos prácticos e ideas para ayudarlo a desarrollarse en su carrera.

¡Suscríbete aquí si eso te suena bien!

IA en cinco | Matt Chapman | Subpila

Las últimas noticias, historias profesionales y consejos de codificación del mundo de la ciencia de datos y la inteligencia artificial, resumidos en 5 viñetas...

aiinfive.substack.com

Fuentes

[1] Oficina de Estadísticas Laborales de EE. UU. Manual de perspectivas laborales: científicos de datos. (8 de septiembre de 2022). <https://www.bls.gov/ooh/math/data-scientists.htm>

[2] Will Stanton. Creando un gran blog de ciencia de datos. (15 de julio de 2015). <https://will-stanton.com/creating-a-great-data-science-resume/>

[3] CV destacado. ¿Cuánto tiempo pasan los reclutadores mirando tu CV? (noviembre de 2022).

<https://standout-cv.com/how-long-recruiters-spend-looking-at-cv#:~:text=Research%20shows%20that%20recruiters%20spend,15%20minutos%20revisando%20a%20CV.>

[4] Jeremy Harris. Las 4 formas más rápidas de no ser contratado como científico de datos. (12 de junio

2018). <https://towardsdatascience.com/the-4-fastest-ways-not-to-get-hired-as-a-data-scientist-565b42bd011e>

[5] Ivanna Kacewica. Configure el sitio web de su cartera en menos de 10 minutos con

Páginas de Github. (1 de marzo de 2019). <https://medium.com/@evanca/set-up-your-portfolio-sitio-web-en-menos-de-10-minutos-con-github-pages-d0efa8ff56fd>

Ciencia de los datos

portafolio

Carreras en ciencia de datos

Carreras

Análisis de los datos



Seguir

Escrito por Matt Chapman

6K seguidores · Escritor para Hacia la ciencia de datos

Científico de datos con sede en Oxford, Reino Unido. Aprenda [SQL en the-sql-gym.com](https://www.sql-gym.com) u obtenga mi plantilla de portafolio en tienda.mattchapman.co

Más de Matt Chapman y Hacia la ciencia de datos



Matt Chapman hacia la ciencia de datos

Deje de abusar de Scikit-Learn y pruebe OR-Tools en su lugar

Muchos científicos de datos abusan del aprendizaje automático y descuidan la optimización matemática, aunque sea Genial para tu carrera y fácil de aprender.



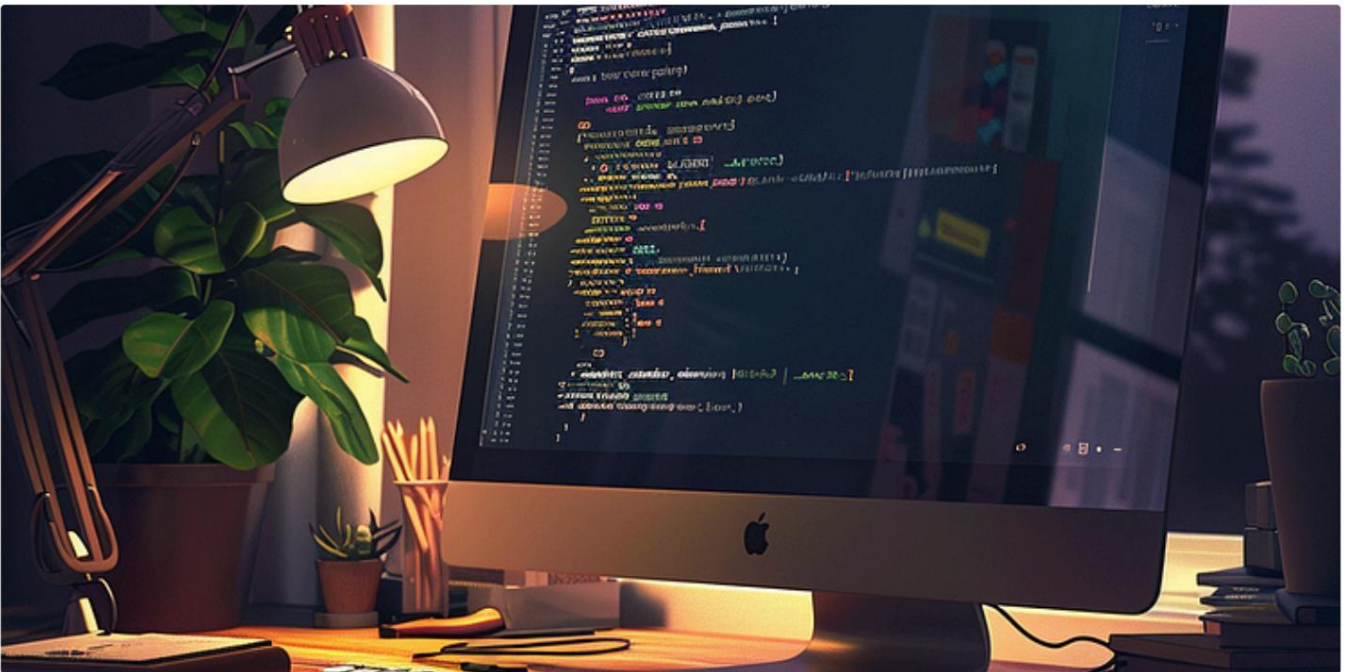
· Lectura de 9 minutos · 26 de enero de 2024



409



6



Krispin Rami en Hacia la ciencia de datos

Configuración de un entorno Python dockerizado: la forma elegante

Esta publicación proporciona una guía paso a paso para configurar un desarrollo dockerizado en Python. entorno con VScode y los Dev Containers...

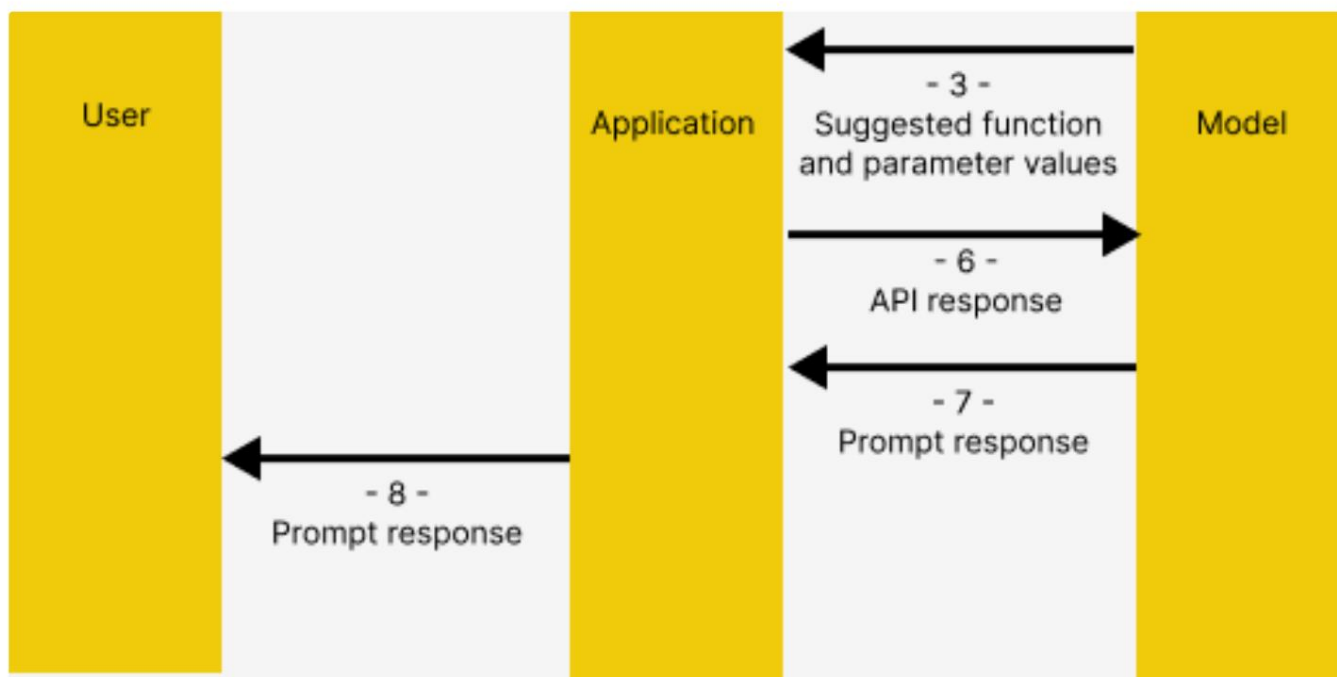
Lectura de 9 minutos 2 de abril de 2024



787



1



Julian Yip hacia la ciencia de datos

Cree agentes de IA autónomos con llamadas a funciones

Transforme su chatbot en un agente que pueda interactuar con API externas

Lectura de 11 minutos 2 de abril de 2024



1K



8





Matt Chapman hacia la ciencia de datos

Cómo me mantengo actualizado con las últimas tendencias en IA como datos de tiempo completo Científico

No, no solo le pido a ChatGPT que me lo diga



· Lectura de 8 minutos · 1 de mayo de 2023



2.8K



31



Ver todo de Matt Chapman

Ver todo de Hacia la ciencia de datos

Recomendado desde Medio

```
*_, a, b, *_ = [1, 2, 3, 4, 5, 6]  
print(_, _)
```

What does this print?

- A) Syntax error
- B) [1] [4, 5, 6]
- C) [1, 2] [5, 6]
- D) [1, 2, 3] [6]
- E) <generator object <genexpr> at 0x1003847c0>



Liu Zuo Lin

Eres decente en Python si puedes responder estas 7 preguntas correctamente

¡¡No hagas trampa por favor!!



· Lectura de 6 minutos · 6 de marzo de 2024



2.7K



19



Natasha Selvaraj  en Hacia la ciencia de datos

Cómo conseguí un trabajo de analista de datos en 6 meses

Y cómo puedes hacer lo mismo en 2023, aprovechando ChatGPT y la IA generativa.

★

Lectura de 10 minutos · 18 de noviembre de 2023

 509

 7





Liza



Modelado predictivo con Python

20 historias · 1117 guardados



Cómo planificar tu carrera cuando ya has comenzado una carrera

10 historias · 327 guardados



Guías prácticas para el aprendizaje automático


10 historias · 1339 guardados



Mensajes de ChatGPT

47 historias · 1456 guardados





Codificación para subir de nivel Somnath Singh

La era de los empleos tecnológicos bien remunerados ha terminado

La muerte de los empleos tecnológicos.


★

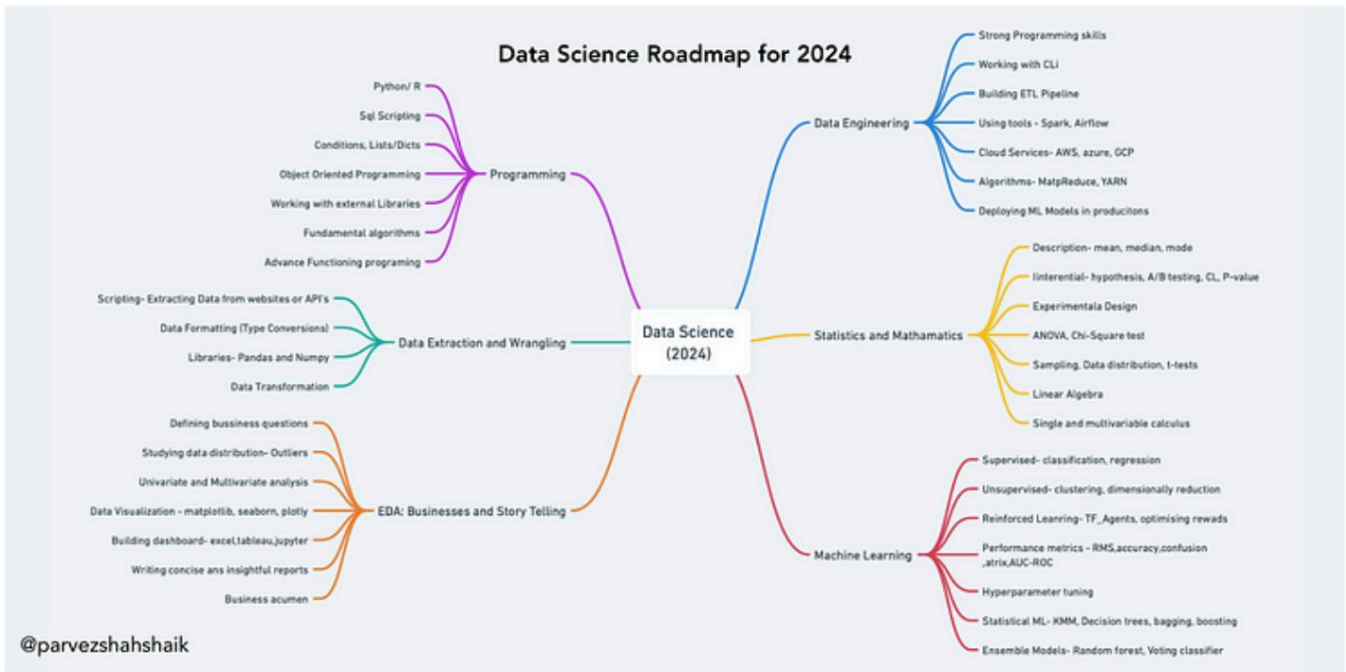
Lectura de 14 minutos · 1 de abril de 2024

 8.5K

 247







Campamento de entrenamiento Parvez Shah Shaik

Hoja de ruta de la ciencia de datos para 2024

Este artículo tiene como objetivo guiarte en la creación de un portafolio sólido que muestre sus habilidades en ciencia de datos. Proporciona un marco, recursos y...

Lectura de 8 minutos 27 de diciembre de 2023



150





Hammad Hassan

Probé Upwork durante 30 días

Me reté a mí mismo a ganar \$100 en el primer mes y resultó que gané más que eso.



· Lectura de 4 minutos · 17 de febrero de 2024



6.1K



125



Paraíso de avellanas

Cómo creo ingresos pasivos sin dinero

Muchas formas de iniciar un ingreso pasivo hoy

Lectura de 5 minutos 27 de marzo de 2024



7.3K



172

[Ver más recomendaciones](#)