En el siguiente texto se informan los resultados de análisis sobre el set de datos de Udemy ([*https://www.kaggle.com/andrewmvd/udemy-courses*](https://www.kaggle.com/andrewmvd/udemy-courses)), para mayor detalle del mismo puede acceder a (url github)

Udemy es una plataforma masiva de cursos abiertos en línea (MOOC) que ofrece cursos gratuitos y pagos. Cualquiera puede crear un curso, este modelo de negocio le permite a Udemy poseer cientos de miles de cursos.

Este conjunto de datos contiene 3.682 registros de distintos cursos, pudiéndose obtener de cada uno de ellos el *título del curso*, la *URL* y *costo* del curso (discriminando a su vez si es gratuito o no).

Otras variables registradas son el *número de subscriptores*, *reviews* y *lecturas*.

También se realizó el registro de la *duración* y *fecha de publicación* de cada uno de los cursos.

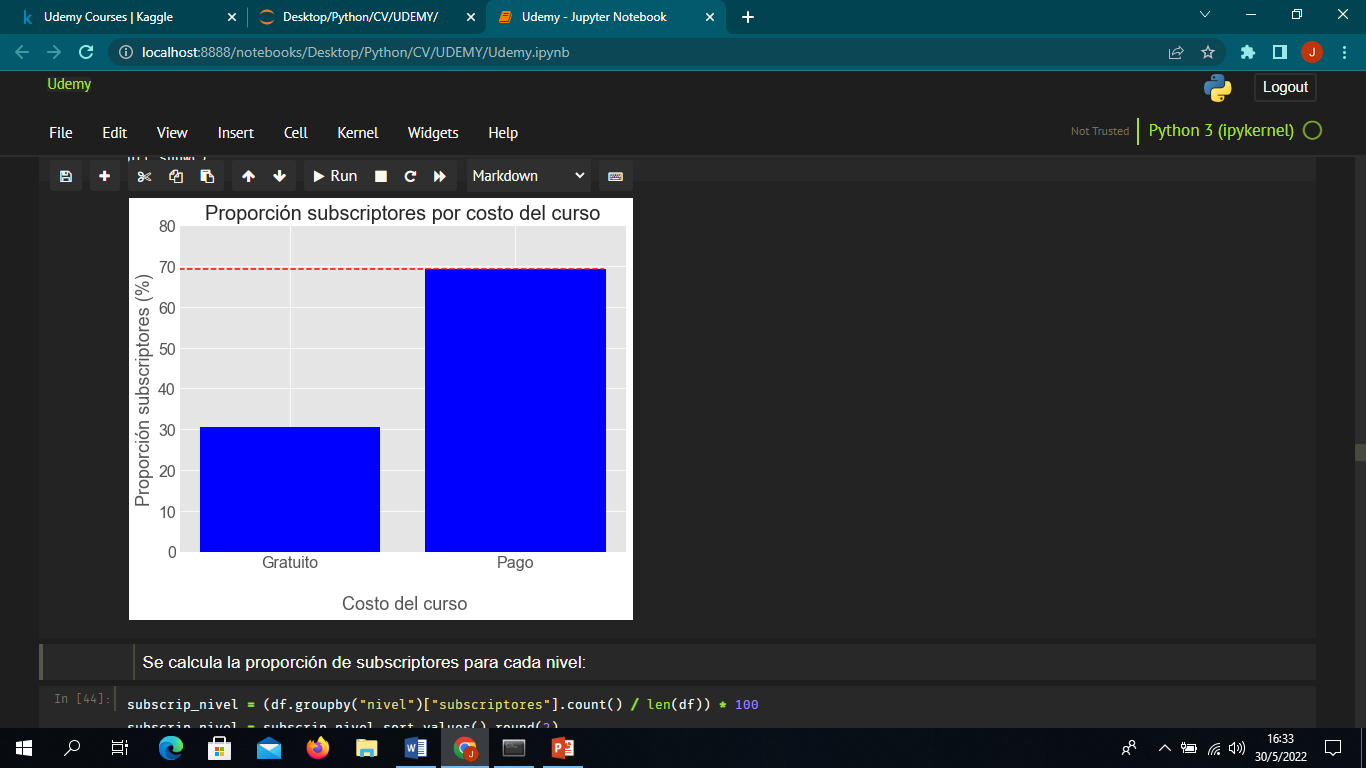
Cada curso se divide en 4 *categorías* distintas:

* Business Finance.
* Graphic Design.
* Musical Instruments.
* Web Development.

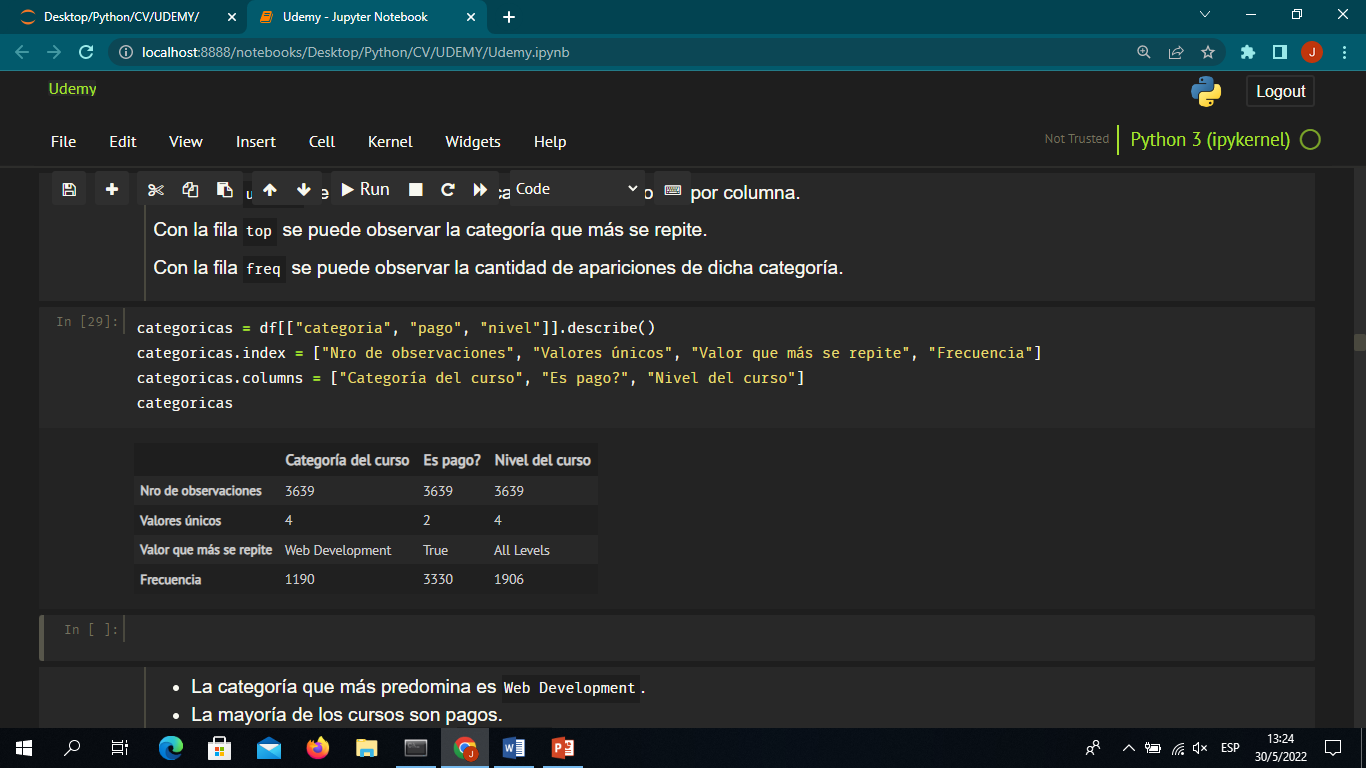
A su vez constan de 4 *niveles* de dificultad:

* All levels.
* Beginner Level.
* Intermediate Level.
* Expert Level.

Del análisis descriptivo del data set se pudo obtener que el 69.38% (*Fig. 1*) de los cursos eran pagos, con una media de precio de 66.16 USD y una mediana de 45 USD, con una duración promedio de 4hs. El 33% de los cursos perteneció a las categorías Web Development y Business Finance, con un nivel orientado a todos los niveles (52%). Del total de cursos, la mayoría perteneció al año 2016 (33%).

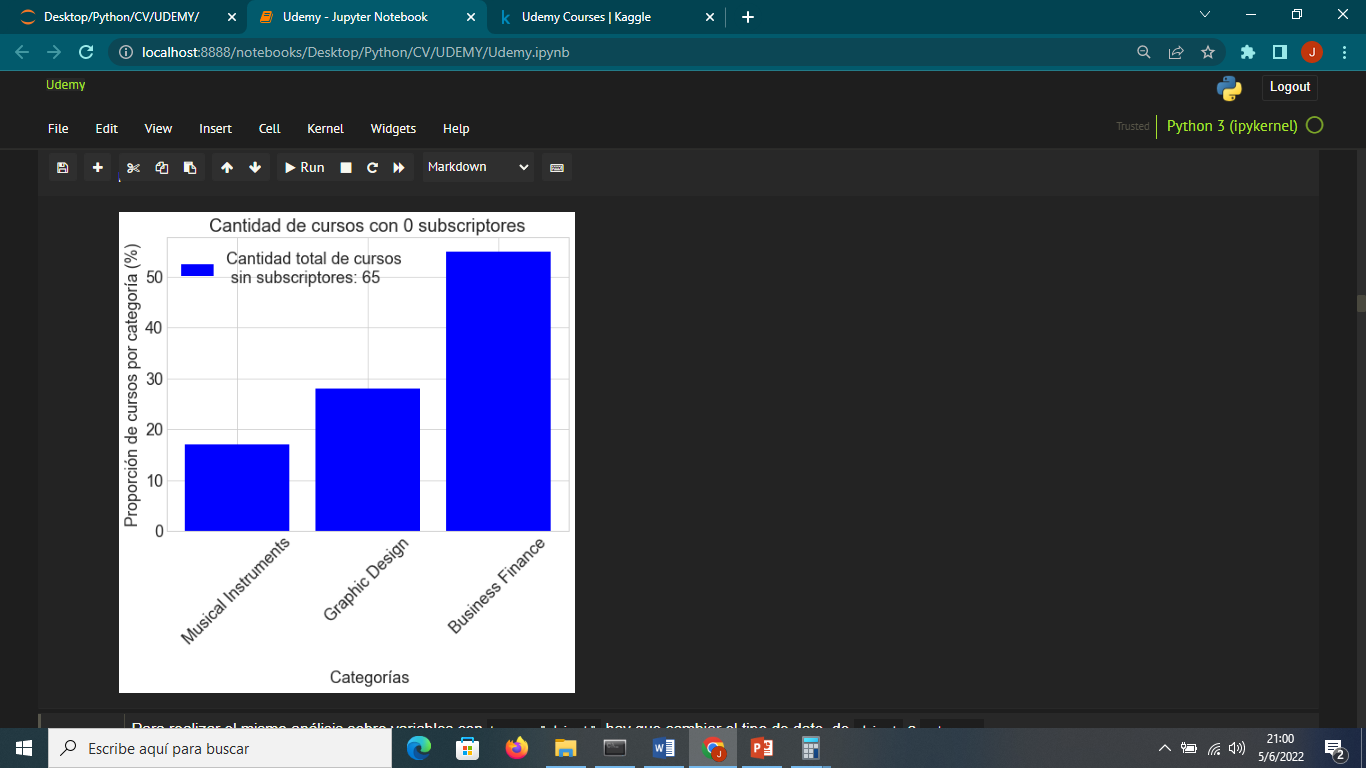


*Fig 1 – Proporción subscriptores por costo del curso.*



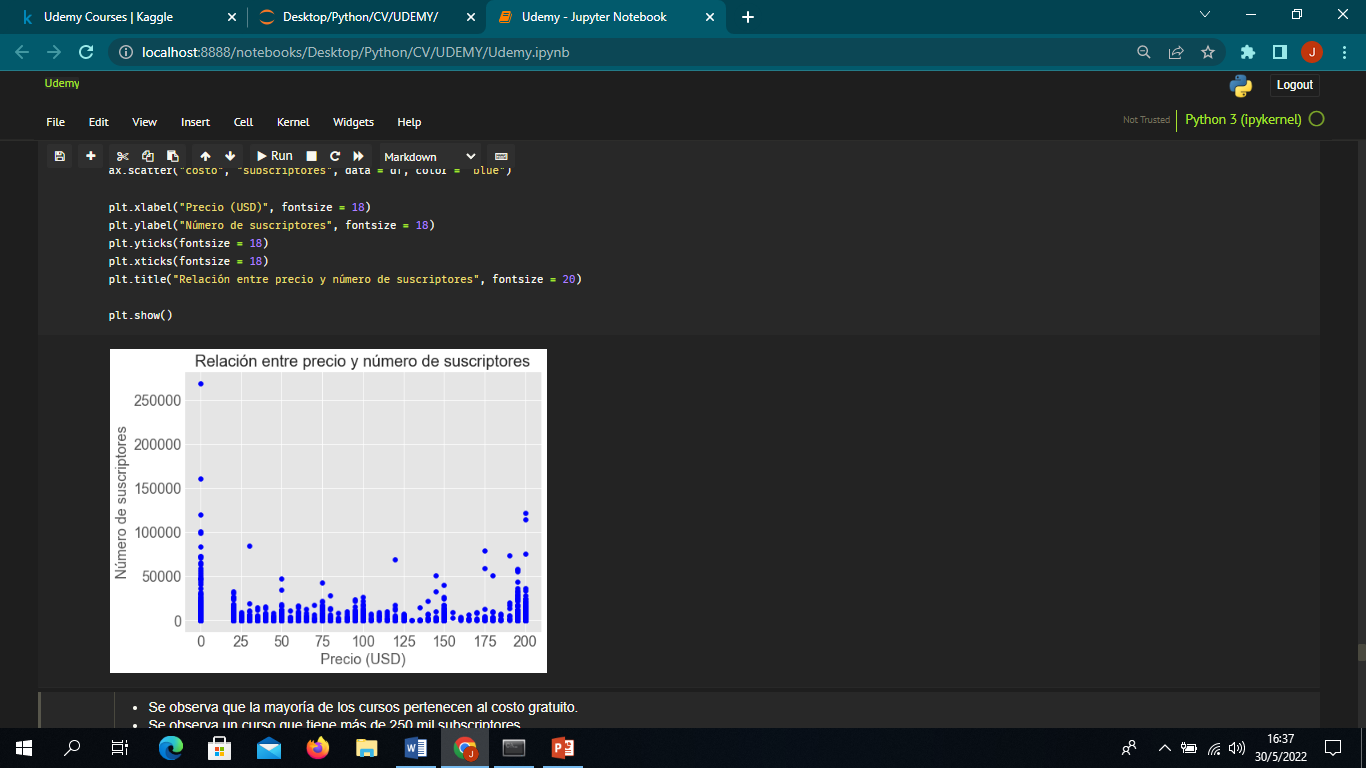
En relación a la Categoría de los cursos

Se observó que una pequeña cantidad de cursos (n =65) no tenían subscriptores (*Fig. 3*). En base a la categoría del curso, se halló que ninguno de ellos pertenecía a Web Development. Más de la mitad de los cursos con 0 subscriptores pertenecía a la categoría Business Finance.

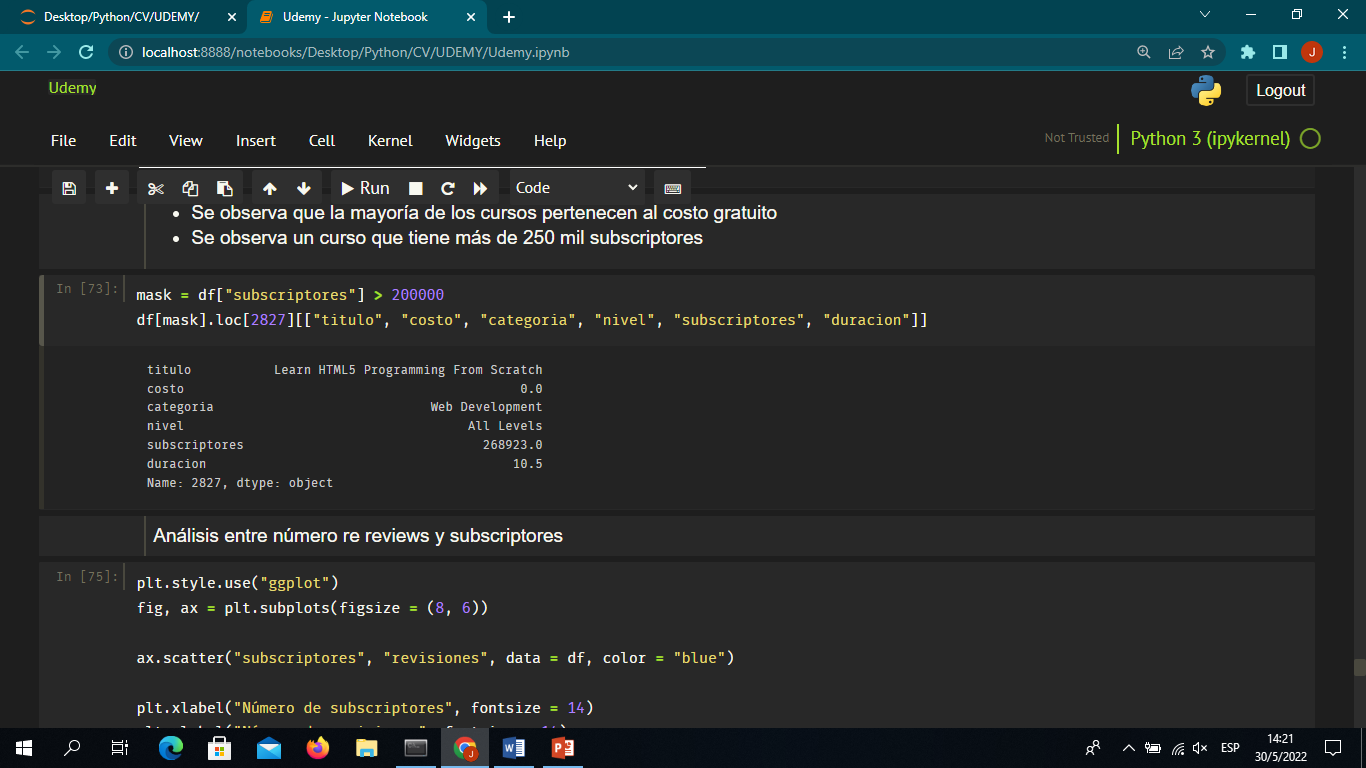


*Fig 3 – Cantidad de cursos con 0 subscriptores.*

Los cursos gratuitos presentaron el mayor registro de subscriptores (n = 309). Llama la atención que existe un curso con más de 250 mil subscripciones (*Fig. 4 y 5*).

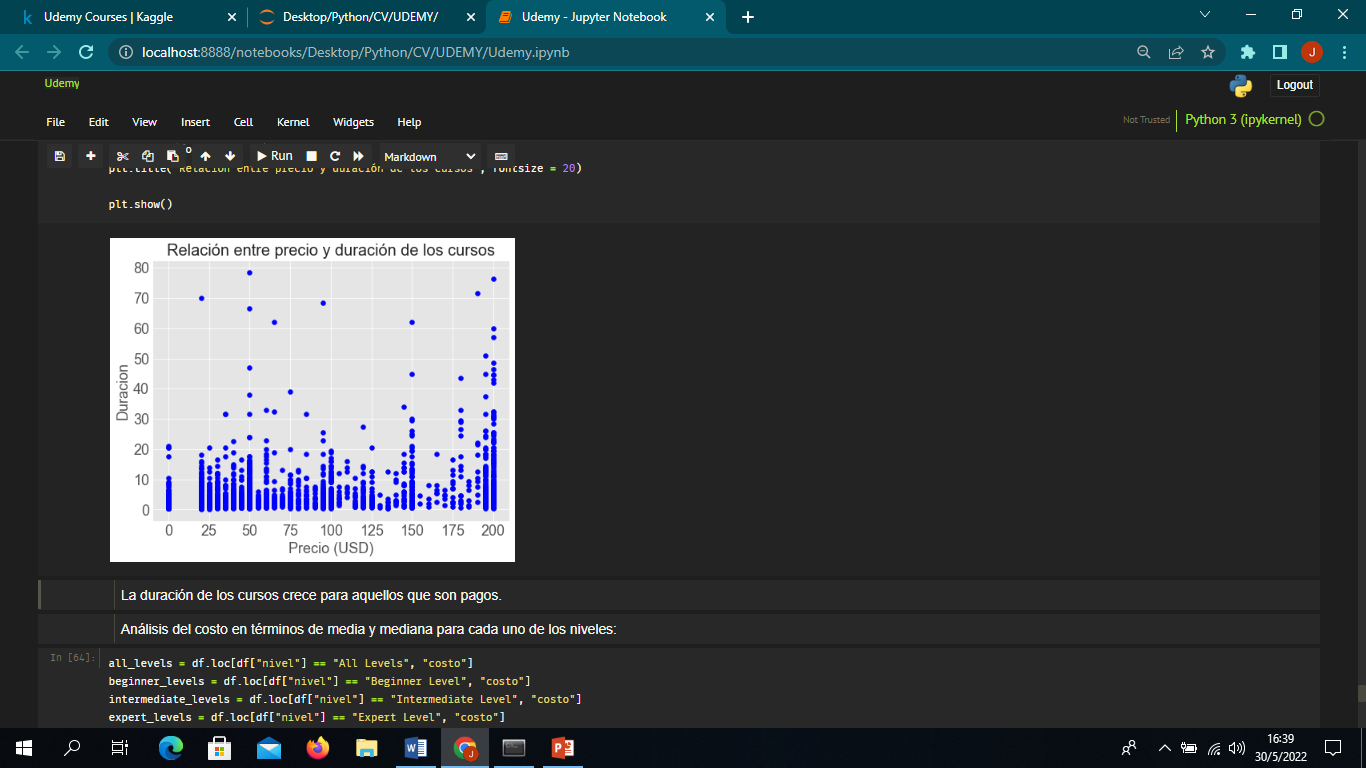


*Fig 4 – Cantidad de subscripciones en base al costo del curso.*



*Fig 5 – Curso que presentó más de 250 subscripciones.*

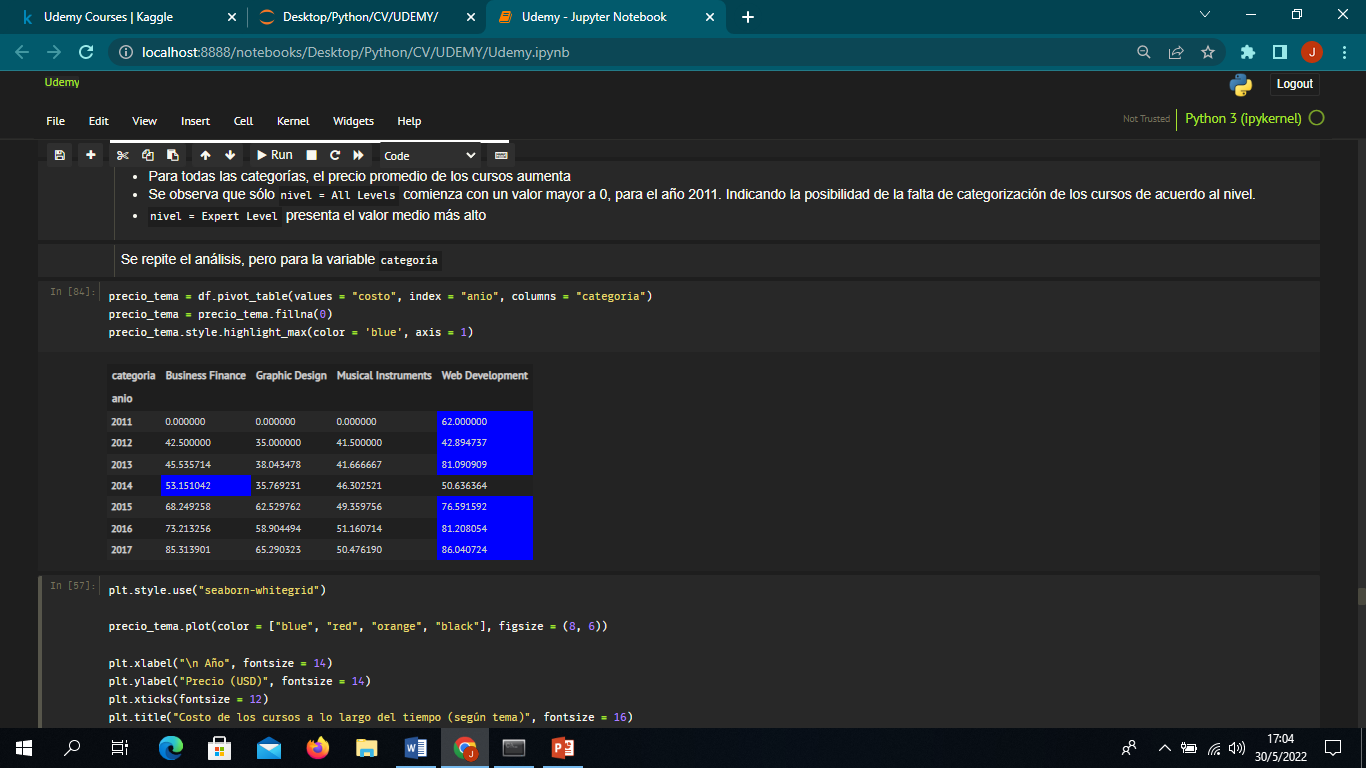
En relación al costo se encontró que los cursos más caros (valores cercanos a 200 USD) presentaban una mayor duración:



*Fig 6 – Relación entre la duración y el precio de los cursos.*

ANALISIS SOBRE FEATURE = CATEGORIA:

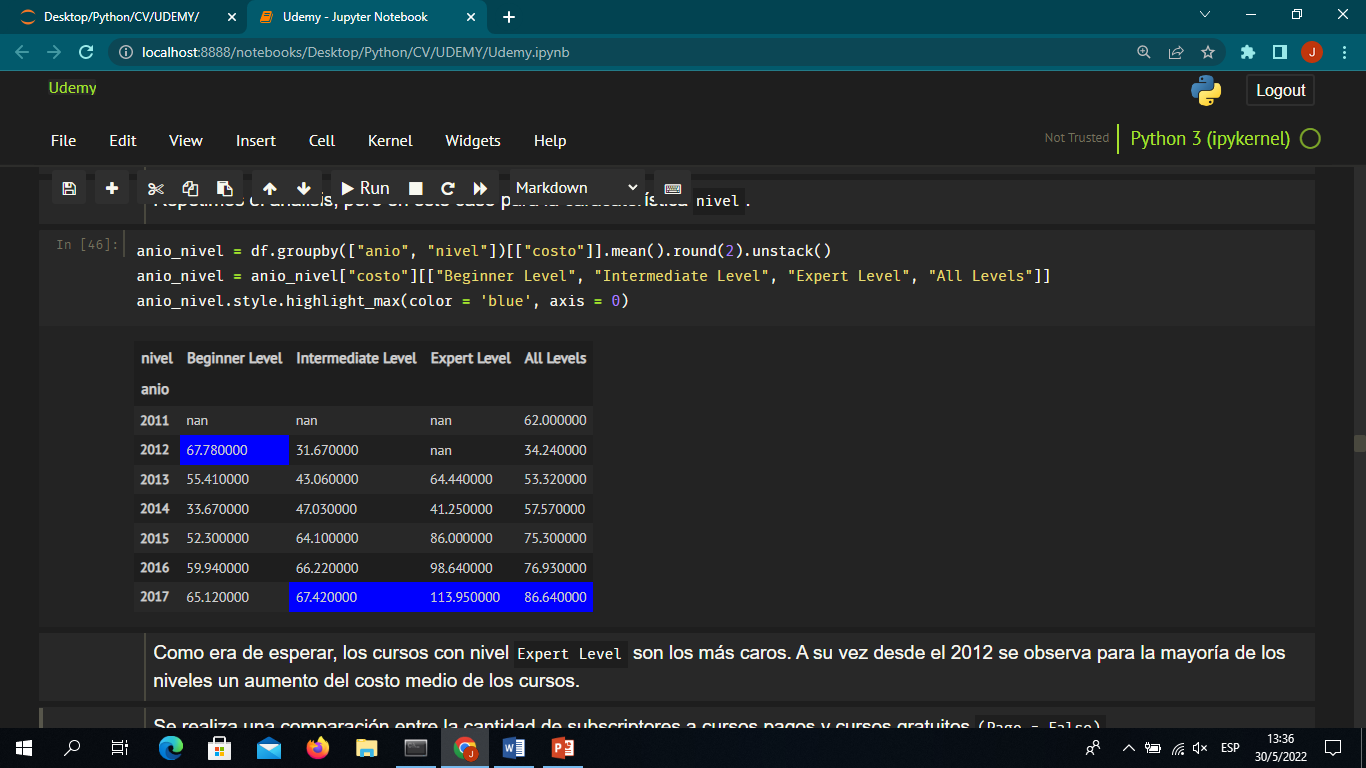
Los cursos de la *categoría Web Development* presentan un costo mayor en promedio para cada año del período 2011-2017.



*Fig 7 – Costo medio de los cursos según la categoría. Resaltado en azul los valores*

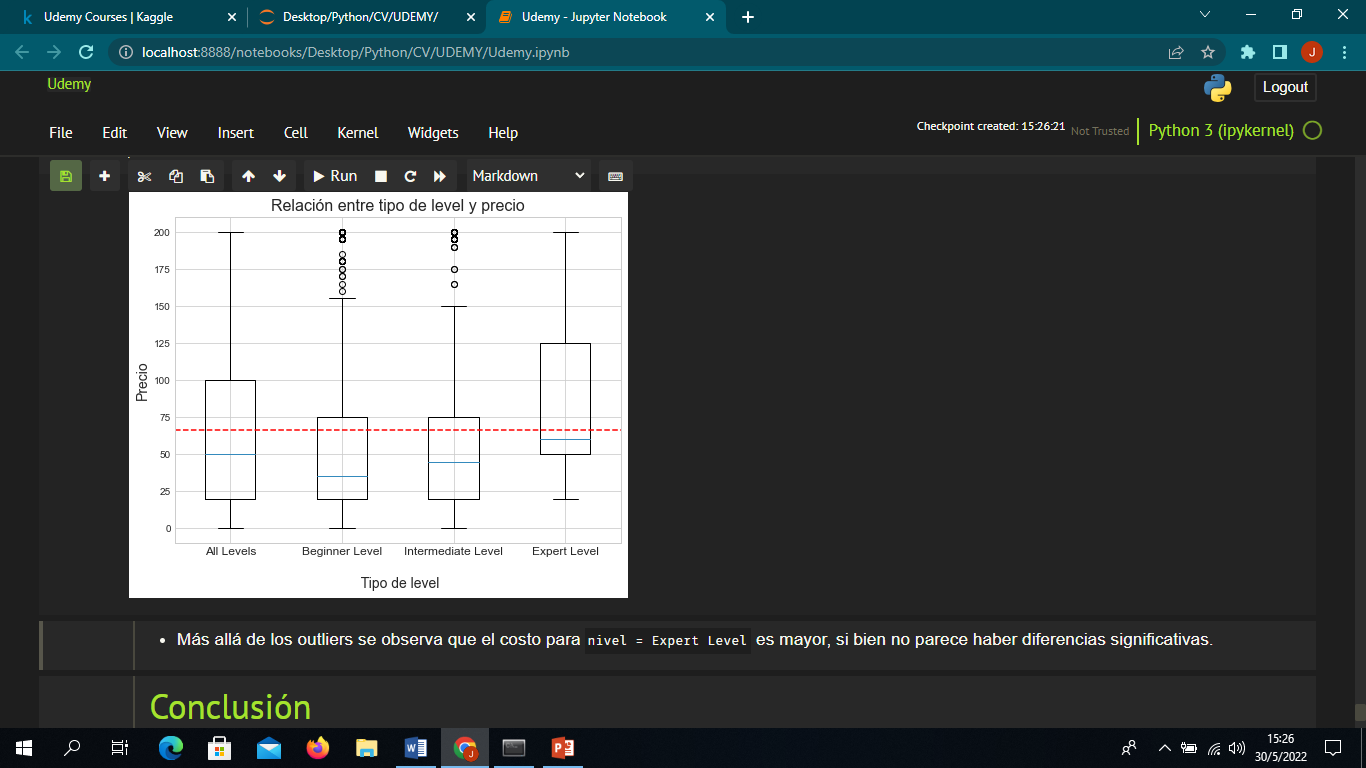
*máximos para cada año.*

Los cursos con nivel Expert Level presentaron una media (*Fig. 7*) y mediana (*Fig. 8*) mayor en cuanto al costo del medio curso, con respecto a los otros niveles.



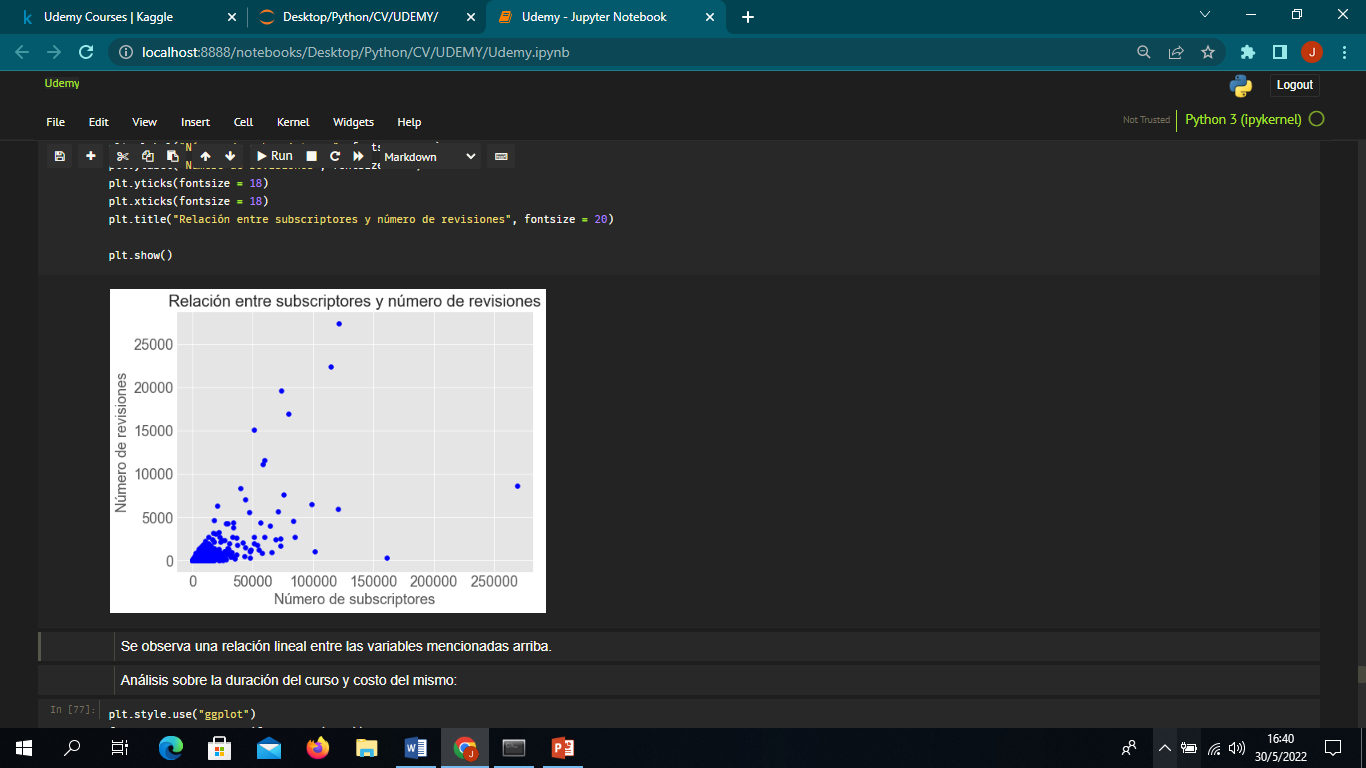
*Fig 7 – Costo medio de los cursos según el nivel. Resaltado en azul los valores*

*máximos para cada nivel.*

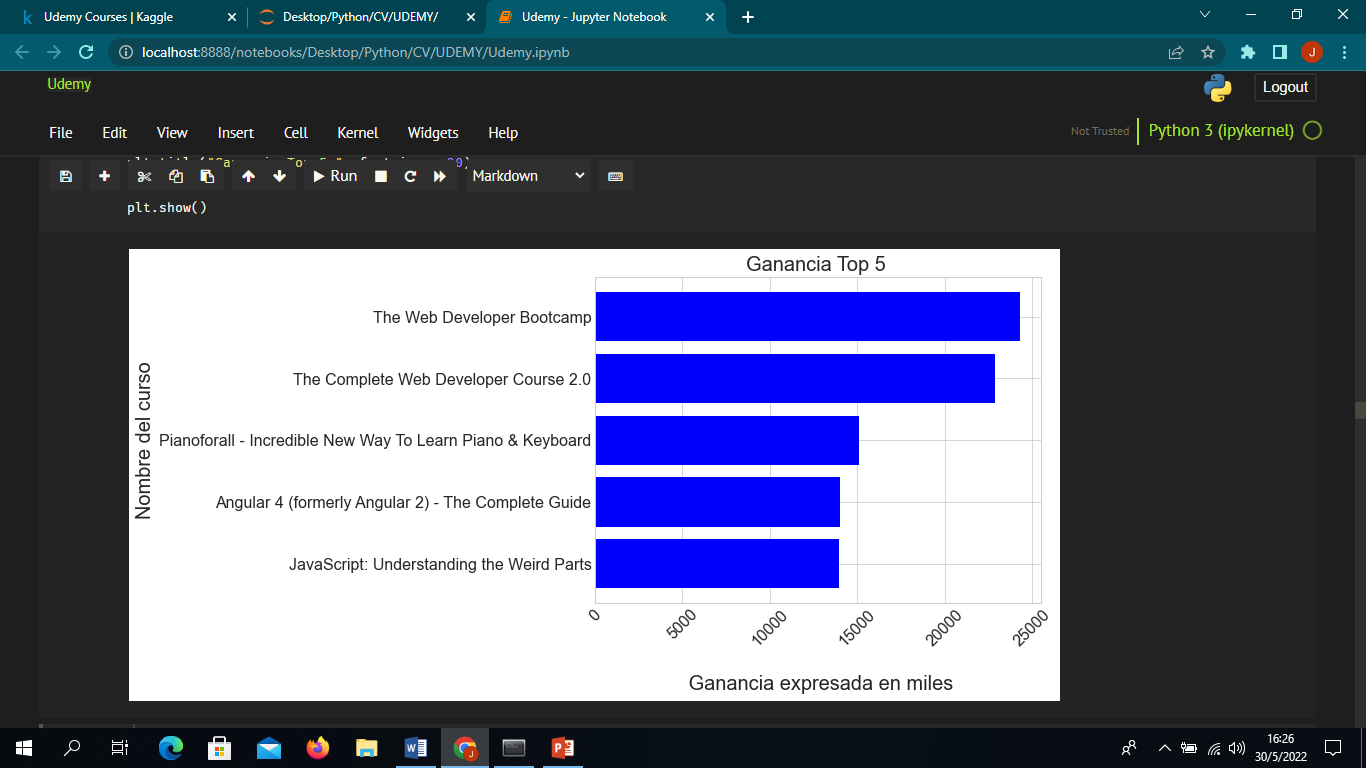


*Fig 8 – Boxplot sobre la relación entre el precio y el nivel de los cursos.*

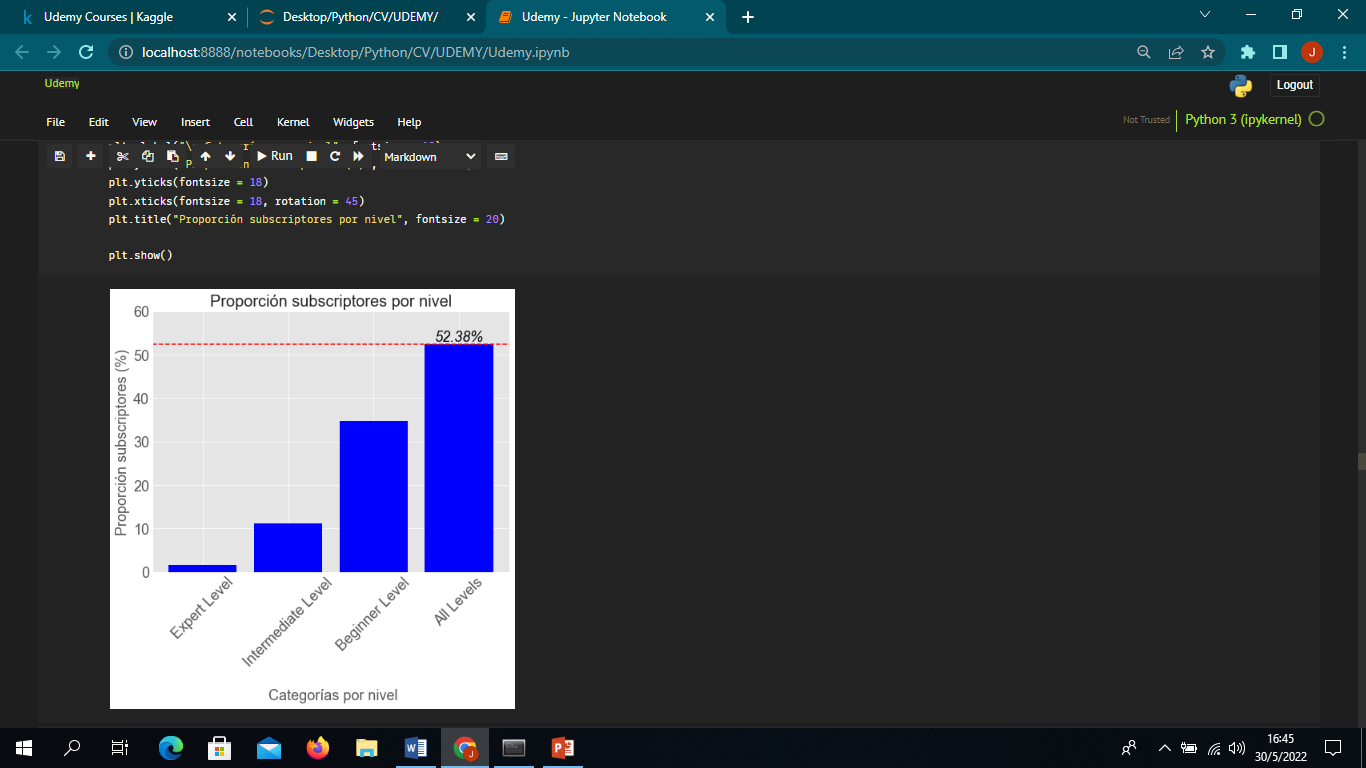
Otro hallazgo fue la relación lineal entre *número de revisiones y número de subscriptores*:

**

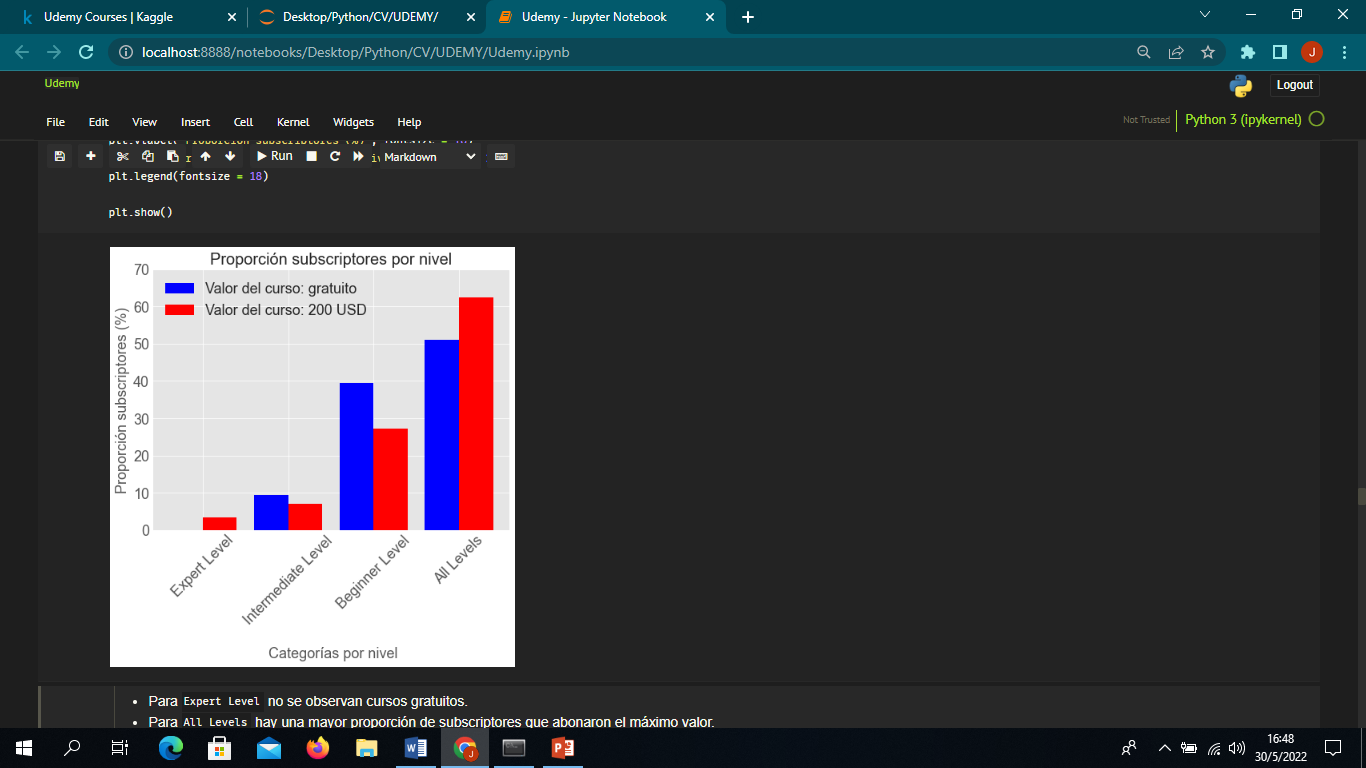
De los 5 cursos que más ganancia dejaron a UDEMY, 4 pertenecen a la *categoría Web Development*.



Respecto a la distribución de los niveles, aproximadamente el 90% de la muestra corresponde a *All Levels y Beginner Level*:

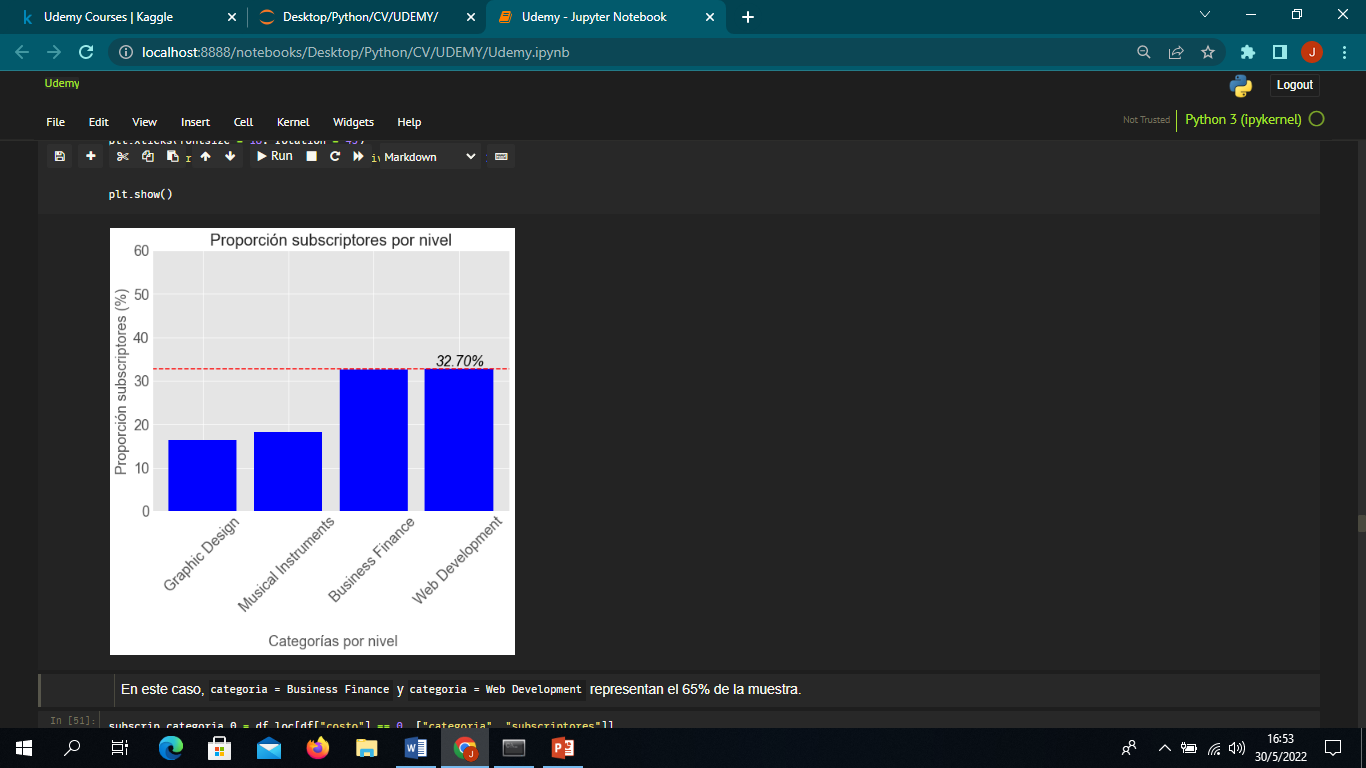


Se profundiza en el análisis, discrimando el *nivel* del curso en base a su *costo (costo = 0 vs costo = 200)*:



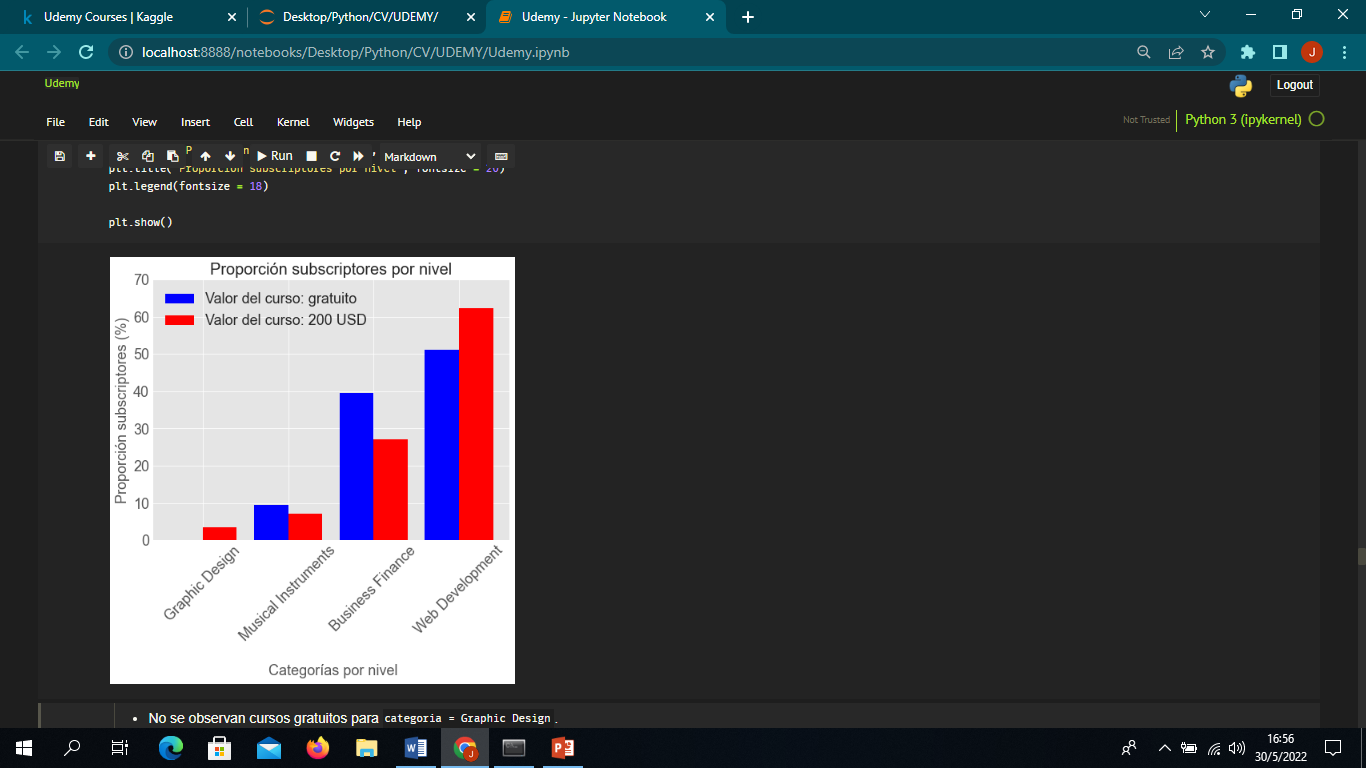
Para *Expert Level* no se observan cursos gratuitos. Respecto a *All Levels*, hay una mayor proporción de subscriptores que abonaron el máximo valor. Finalmente, *Beginner Level* invierte la tendencia recién mencionada, siendo mayor la proporción de subcriptores para cursos gratuitos.

Se repite el análisis, pero para la variable *categoría*. Luego se compara dentro de cada categoría la proporción de subscriptores con *costo = 0* y *costo = 200.*



En este caso, las *categorías Business Finance* y *Web Development* representan el 65% de la muestra.

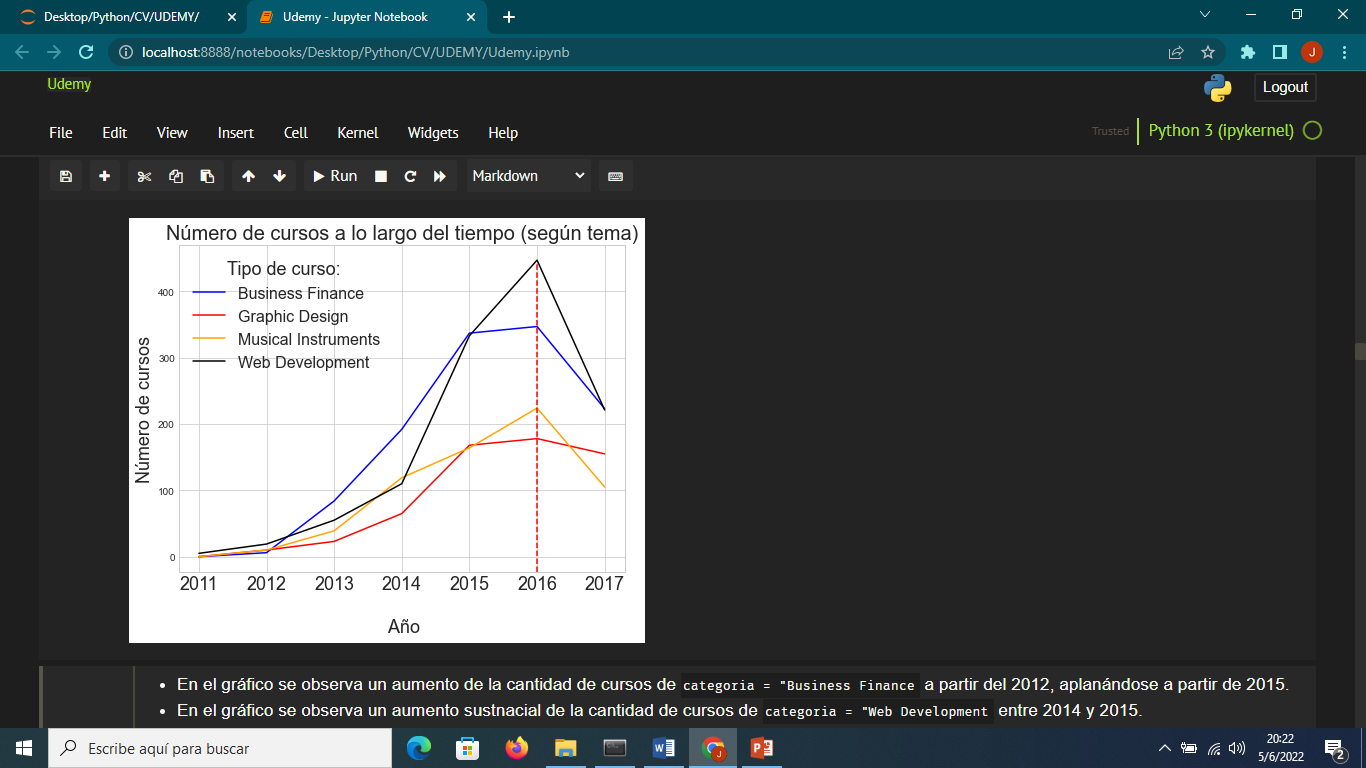
Al discriminar respecto al *costo* del curso (*cost = 0*  y *costo = 200*) se observa que no hay curos gratuitos para *Graphic Design*. Los cursos de *Web Development* tienen una mayor proporción de subscriptos que abonaron el precio de 200 USD, relación que se invierte con los cursos de *Business Finance.*



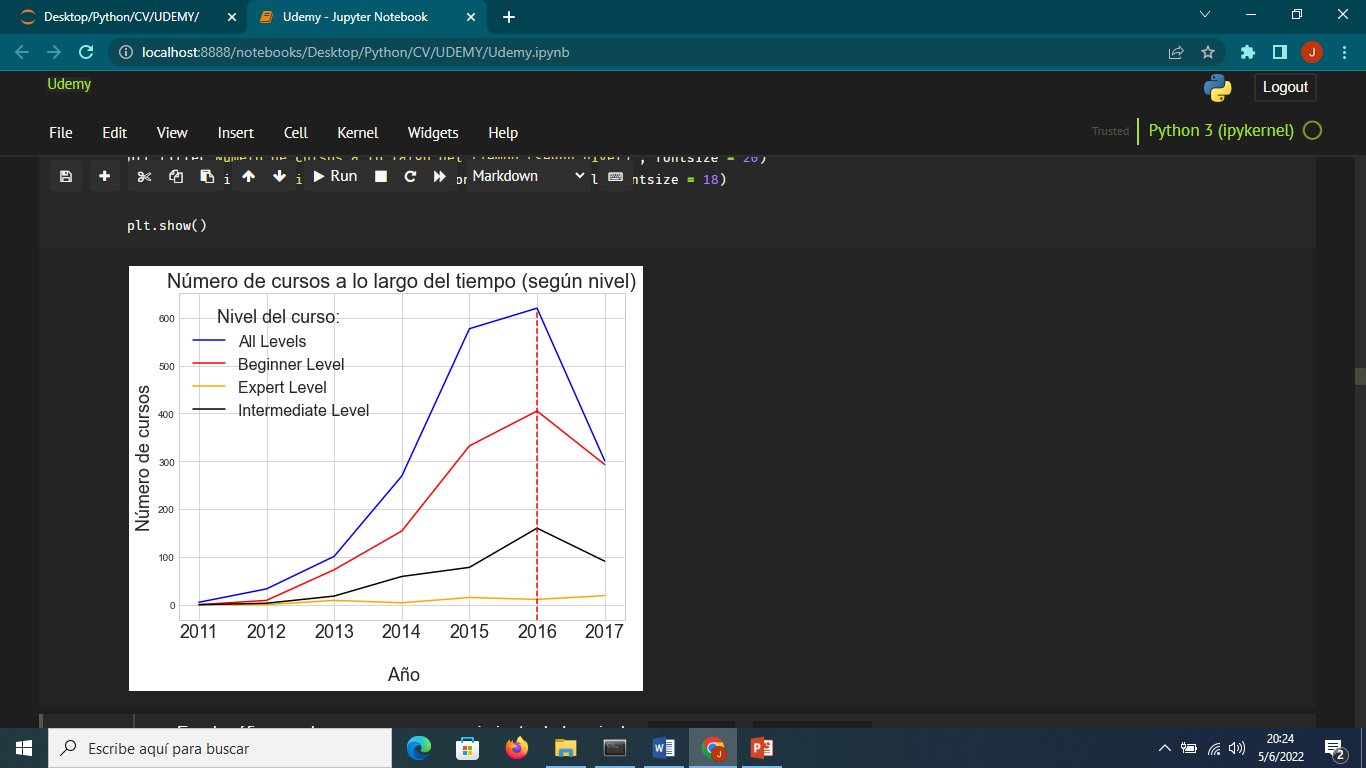
Teniendo en cuenta la variable “tiempo”, podemos realizar 3 tipos de análisis diferentes:

* Relación de la cantidad de cursos a través del tiempo (discriminando por categoría o nivel).
* Relación del precio de cursos a través del tiempo (discriminando por categoría o nivel).
* Relación de la cantidad de subscriptores a través del tiempo (discriminando por categoría).

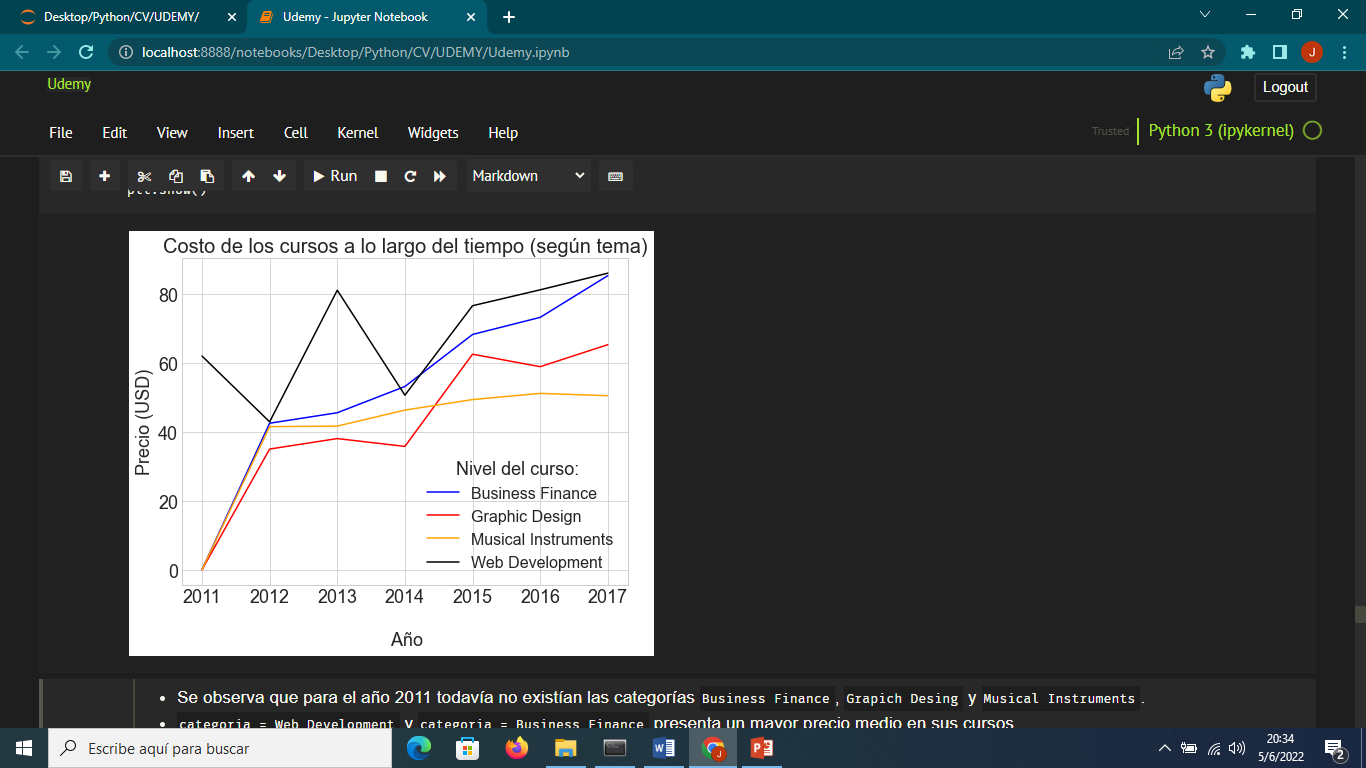
En el siguiente gráfico se puede observar que los cursos de Udemy a principios del 2011 no estaban discriminados en base a su categoría. A su vez las categorías presentan un crecimiento en relación a la cantidad de cursos, con una caída marcada a partir del año 2016. En dicho año, la categoría Business Finance tenía la mayor cantidad de cursos.



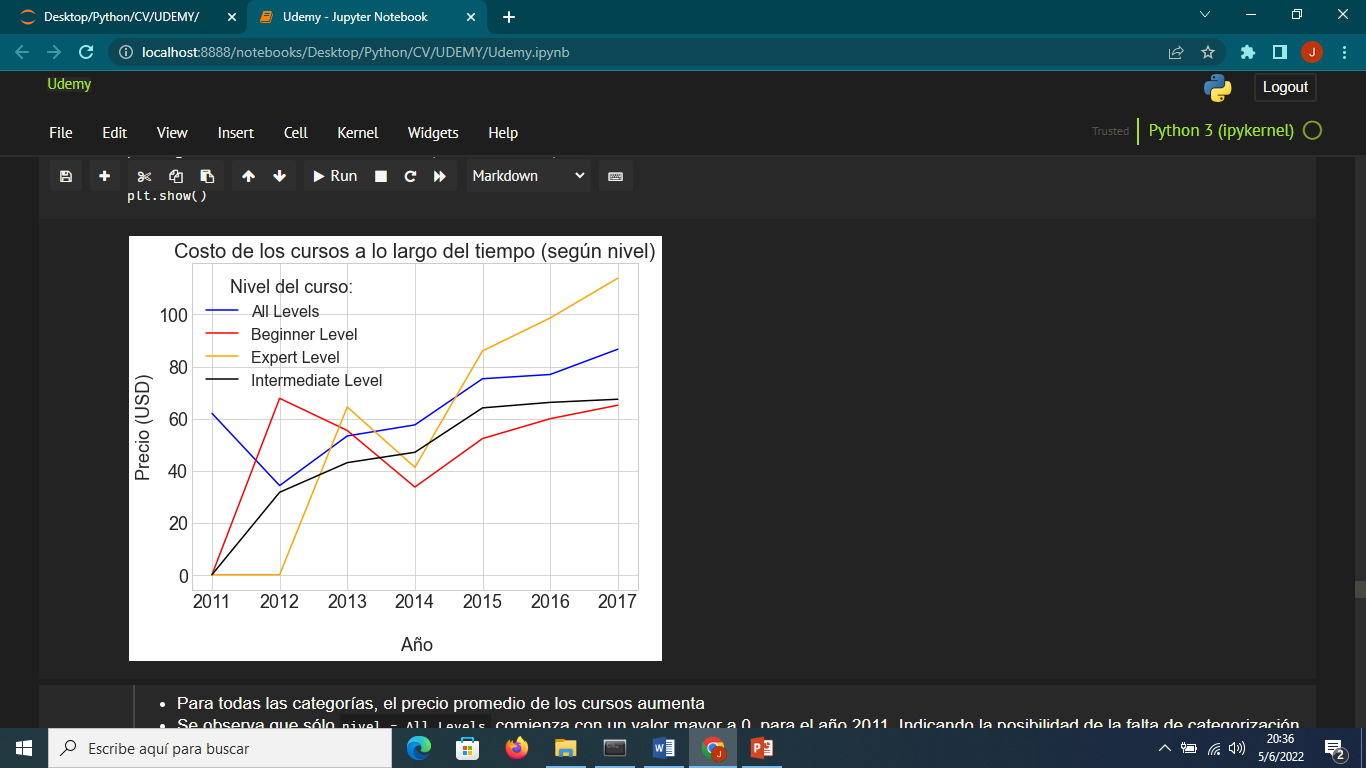
Una situación similar se observa en cuanto a los niveles de los cursos. A principios del 2011 los cursos de Udemy no estaban discriminados por nivel, a partir de dicho año casi todas las categorías de nivel crecen con excepción de los cursos de nivel experto. Nuevamente aparece un pico en el 2016, a partir del cual el descenso en la cantidad de cursos es muy marcado.



Según lo indicado en el gráfico, a principios del 2011 Udemy los cursos de Web Development rondaban por los 60 dólares. En el 2013 hubo un pico, acompañado de un descenso en el año siguiente. El resto de las categorías aumentó su precio medio a lo largo del período 2011-2017. Business Finance y Web Development son los cursos más caros.



En relación al costo medio de los cursos según el nivel del mismo se observa un aumento, siendo los cursos de nivel Experto Level lo más caros.



* Todos los cursos para diferentes categoria presentan una caída en el número de subscriptores
* Se observa que para las categorias = Business Finance, categorias = Graphic Design, categorias = Musical Instruments no hay subscriptores para el año 2011.
* Las bajas más marcadas para las categorias Musical Instruments y Graphic Design se dan a partir del año 2012 - 2013
* Para la categoria = Web Development a partir del año 2013 - 2017

