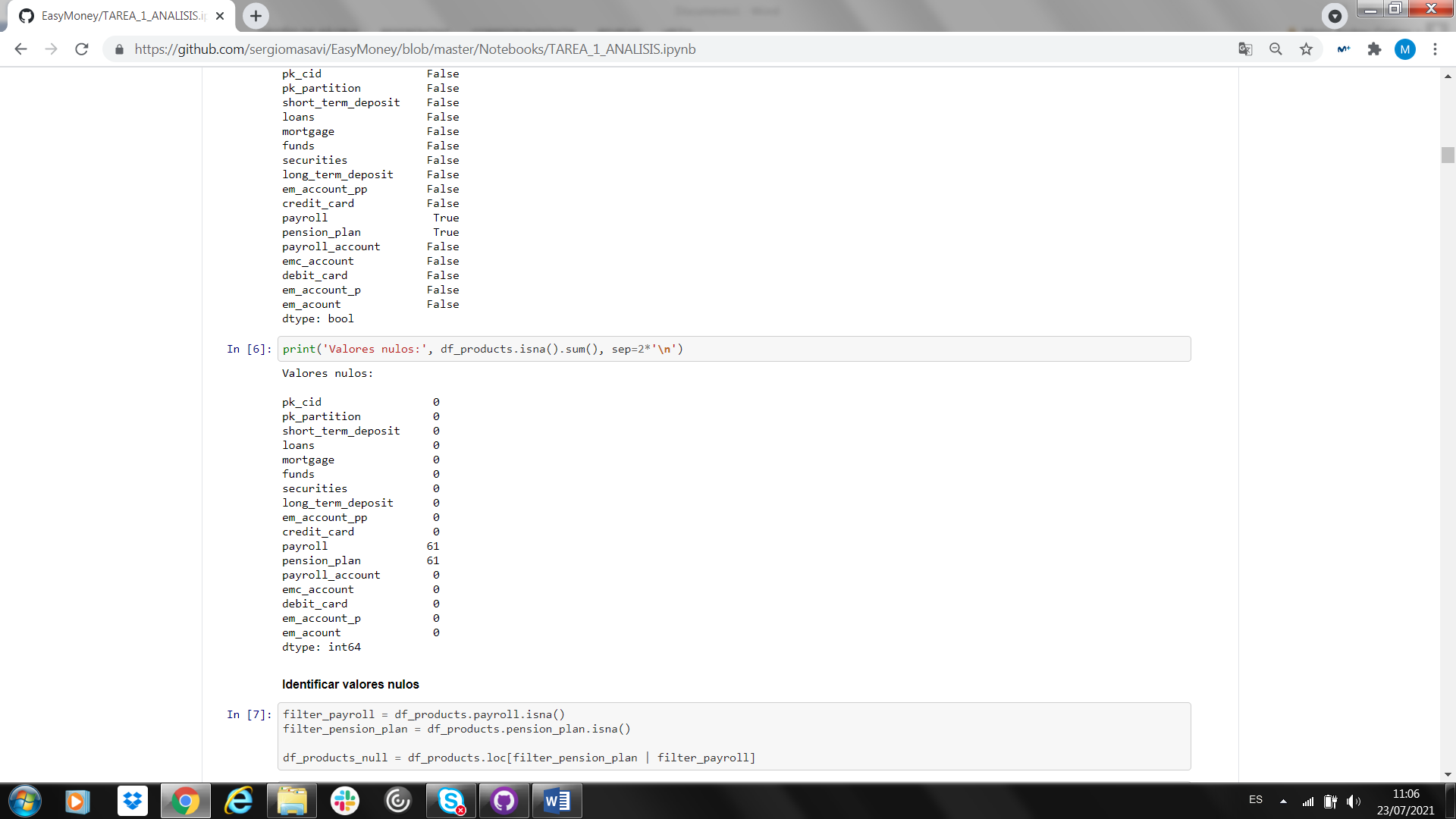
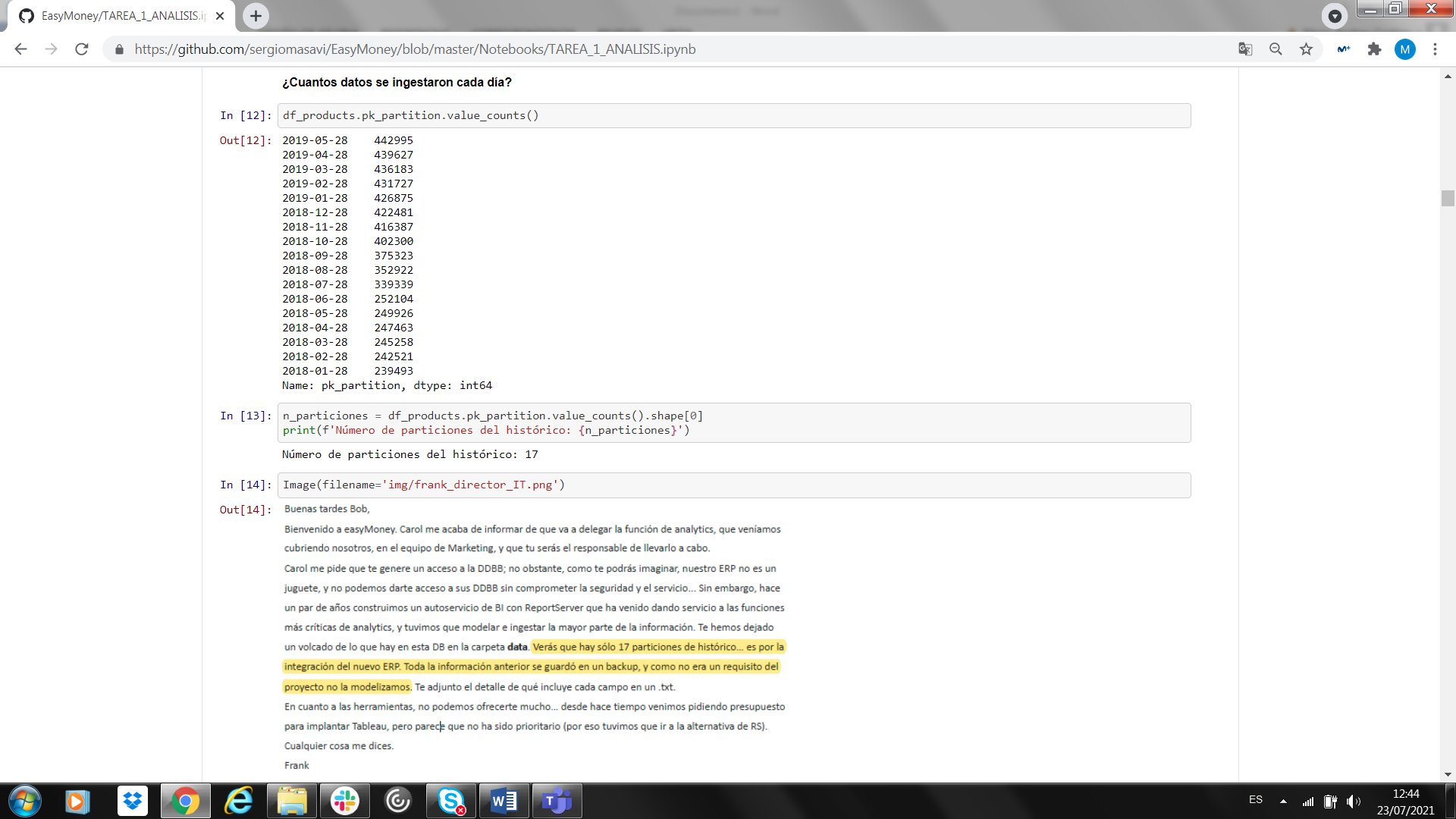
**ANALISIS DE DATOS**

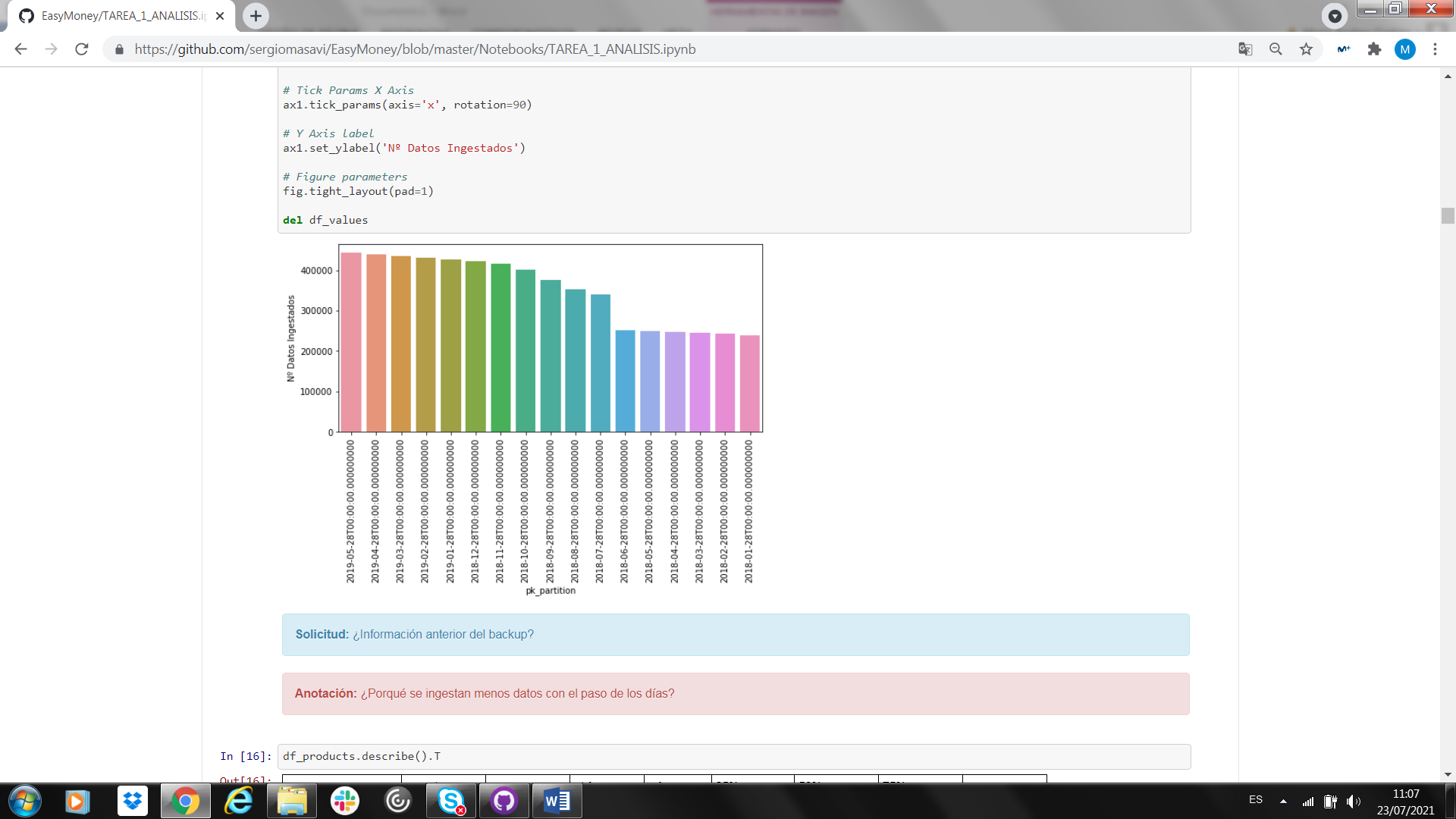
Nos encontramos analizando el Dataset “Products”, y se nos plantean una serie de interrogantes tras haber analizado cada varaible de manera independiente y habiendo conocido las características generales del mismo (índice, forma, columnas, registros etc…):

1. **¿Cuál son los requisitos para ser cliente de EasyMoney?**
   1. Es preciso tener una cuenta básica Easy Money?
   2. Se puede escoger el servicio que se desee?
2. **Que significan los 61 valores nulos?:** (payroll y pension\_plan)
   1. No se han introducido los datos?
   2. No se ha contratado ese servicio?  
      
3. Entendemos que la ingesta de datos, es el volcado de los datos mensuales, pero ¿**es posible saber la fecha de apertura de contratación del servicio del cliente?**
4. **Vemos que cada mes, la cantidad de datos aumenta, ¿a qué es debido?**
   1. Nuevos servicios contratados, incluyendo los nuevos clientes? 

|  |  |  |
| --- | --- | --- |
| **Fecha volcado de datos** | **cantidad de datos volcados** | **Dif mes a mes** |
| 2018-01-28 239493 | 239493 |  |
| 2018-02-28 242521 | 242521 | 3028 |
| 2018-03-28 245258 | 245258 | 2737 |
| 2018-04-28 247463 | 247463 | 2205 |
| 2018-05-28 249926 | 249926 | 2463 |
| 2018-06-28 252104 | 252104 | 2178 |
| 2018-07-28 339339 | 339339 | 87235 |
| 2018-08-28 352922 | 352922 | 13583 |
| 2018-09-28 375323 | 375323 | 22401 |
| 2018-10-28 402300 | 402300 | 26977 |
| 2018-11-28 416387 | 416387 | 14087 |
| 2018-12-28 422481 | 422481 | 6094 |
| 2019-01-28 426875 | 426875 | 4394 |
| 2019-02-28 431727 | 431727 | 4852 |
| 2019-03-28 436183 | 436183 | 4456 |
| 2019-04-28 439627 | 439627 | 3444 |
| 2019-05-28 442995 | 442995 | 3368 |
|  | TOTAL | 203502 |

Hay un incremento de datos desde el 28 de enero del 2018 hasta el 28 de mayo de 2019, (que es el único registro de 203.502 datos. Siempre con una tendencia ascendente. Observando que en fechas como julio del 2018, se produce casi el 25% de los volcados totales. **¿ a que es debido?** Y los meses de septiembre y octubre, el volumen de datos también es importante.

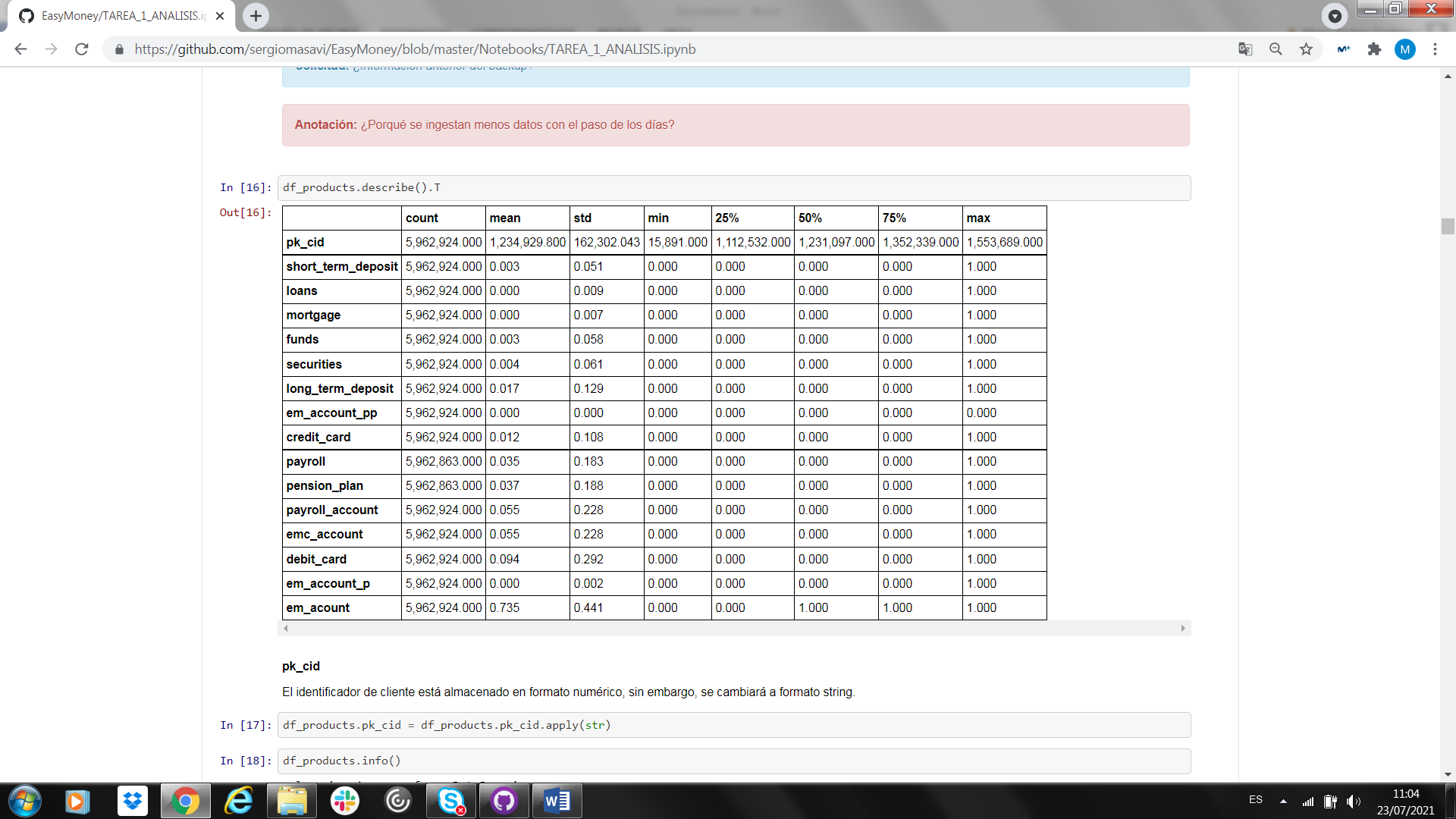
(se debe a alguna campana de mk, (descuentos, oferta de captación) nueva apertura de servicios? Nuevas condiciones?....

* 1. Observamos que NO se reflejan las bajas de los clientes. De qué manera obtenerlas? Averiguar cuando dejan de volcarse datos de esos clientes…

1. Si en un mes, no hay descarga de datos de un cliente, **es porque se ha dado de baja? O porque no ha contratado ningún servicio?**
2. Hemos encontrado, algunos clientes, que solo tienen contratado algunos meses la tarjeta de debito, **¿no debería ir asociada una tarjeta a una cuenta?**
3. Por otro lado, nos encontramos que tenemos en cartera 456.373 clientes, de los cuales………… no tienen ningún servicio contratado. Significa eso, que **¿no son clientes reales? Y son potenciales clientes?**

**NUEVOS PLANTEAMIENTOS.**

1. ¿**Podríamos saber, cuantos clientes tienen contratado un único servicio, 2,3…. Servicios? ¿Cuál es el cliente con más servicios contratados?**



El “Describe” nos indica que:

* El 75% de los clientes tiene un servicio contratado que es la cuenta básica de easy money.
* Observamos que la media de:
  1. Prestamos ( loans) es 0
  2. Hipotecas (Motgage) es 0
  3. Cuenta Premium (em:account\_p es 0

Pero que en su std si recibe un %, y en el max, contienen un 1**. No consigo cuadrar estos datos.**

1. **Según el std y cotejado con la media;**

* Los servicios **MÁS CONTRATADOS:**

**Std mean**

* 1. Em\_account 0,441 0,735
  2. Debit card 0,292 0,094
  3. Payroll\_account 0,228 0,055

emc\_account

* Los servicios **MENOS CONTRATADOS;**

**Std mean**

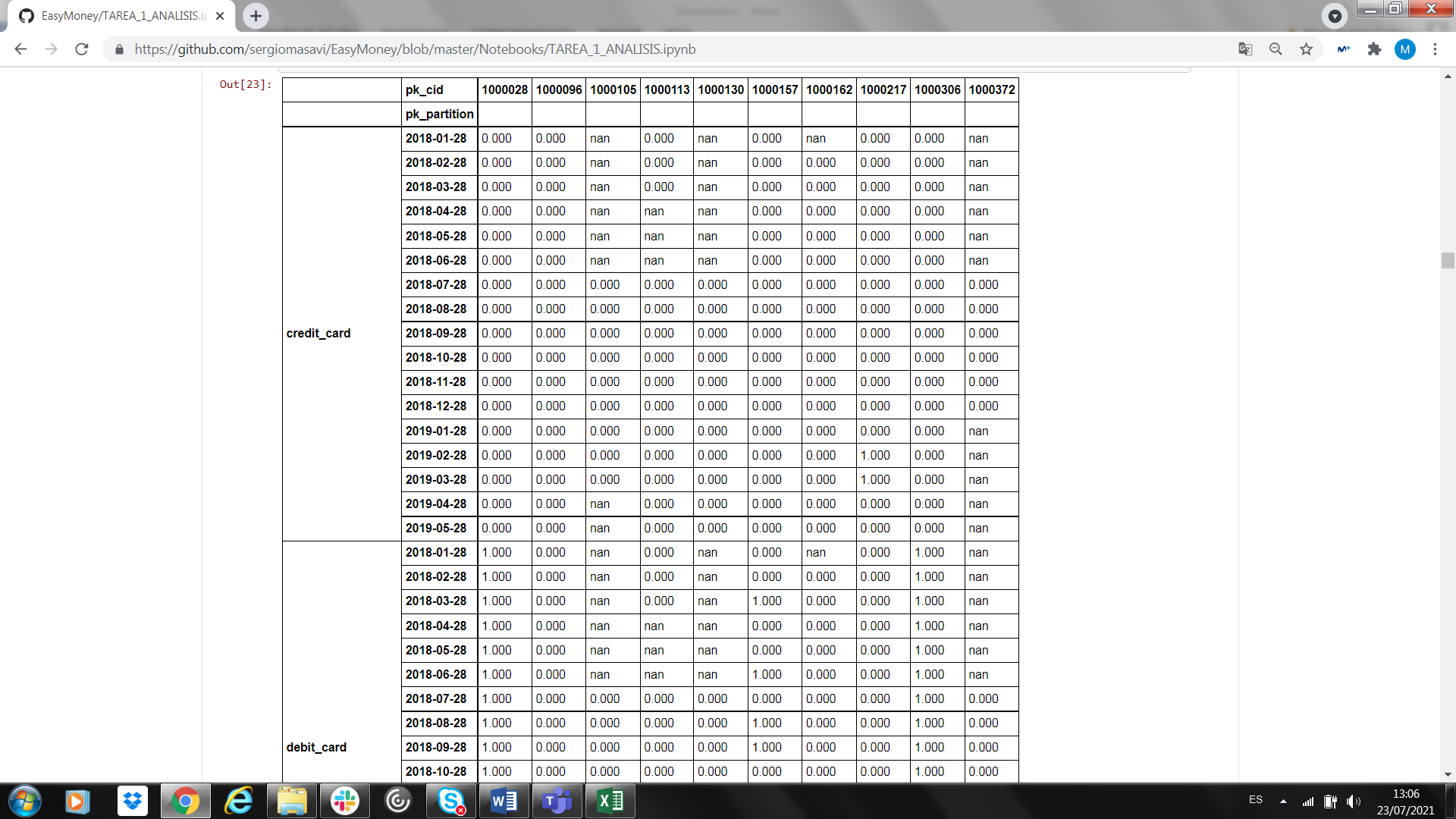
* 1. Prestamos 0,009 0
  2. Hipotecas 0,007 0
  3. Depósitos a c/p 0,051 0,003

Fondos

1. **% de fidelización de clientes**?

* Cuantos clientes están desde el inicio?
* Cuantos clientes se han dado de baja?
* Cuantos hemos recuperado?

1. ¿Por qué aparecen nulos en registros de otros servicios y que, cuando cotejamos los nulos, nos mostró que esos servicios no tenían datos nulos?



Una vez empecemos analizar el dataset socio demographic, deberemos combinar “products & sociodemographic” y analizar:

* De donde son los clientes que compran nuestro servicio y por que.
* Que publico objetivo tiene cada servicio
* Que rango de edad contrata cada servicio
* Son hombres o mujeres
* Con que media salarial