



**Data Glacier**

Your Deep Learning Partner

# Business and Data Understanding

Week 7 - LISUM 04

**José Vicente Solorzano**

**11/08/21**

# Agenda

Data Understanding

Clients

Last contact

Campaign

Bivarial analysis of some features

Recommendations

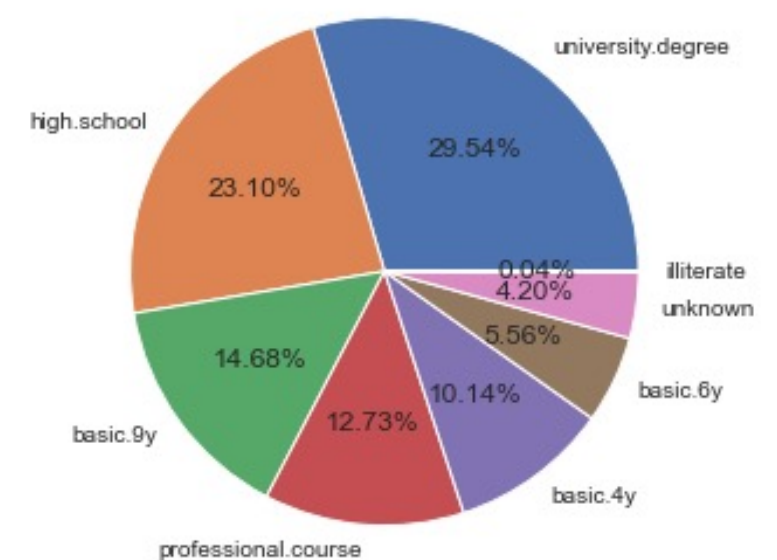
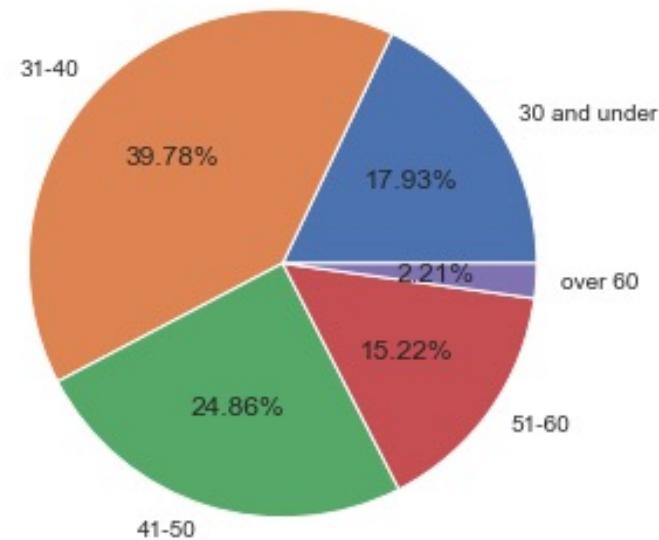
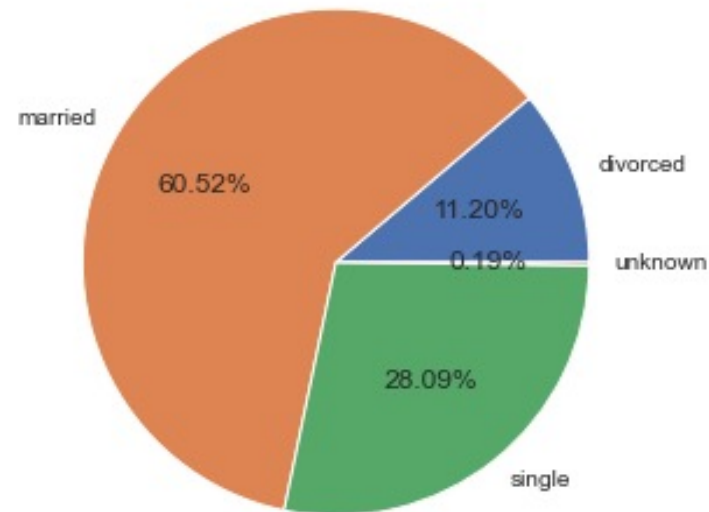
Features are divided into 4 groups:

1. Data related to clients (age, job, marital, education, default, housing and loan).
2. Data related with the last contact of the current campaign (contact, month, day\_of\_week and duration).
3. Data related with campaigns (campaign, pdays, previous and poutcome).
4. Data related with socio economic context attributes (emp.var.rate, cons.Price.idx, cons.conf.idx, euribor3m, nr.employed).
5. TARGET: “y”.

---

Number of rows: 41188 – Number of columns: 21

```
['age', 'job', 'marital', 'education', 'default', 'housing', 'loan',  
'contact', 'month', 'day_of_week', 'duration', 'campaign', 'pdays',  
'previous', 'poutcome', 'emp.var.rate', 'cons.price.idx',  
'cons.conf.idx', 'euribor3m', 'nr.employed', 'y'],
```



### Customers job (%):

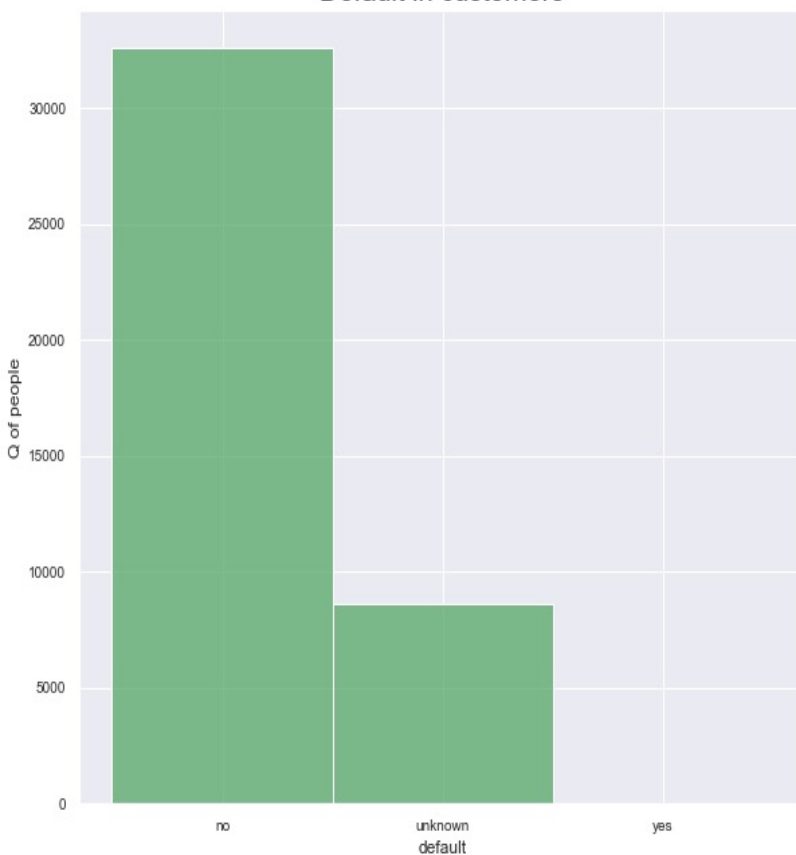
job	
admin.	25.303486
blue-collar	22.467709
technician	16.371273
services	9.636302
management	7.099155

### Customer's demo:

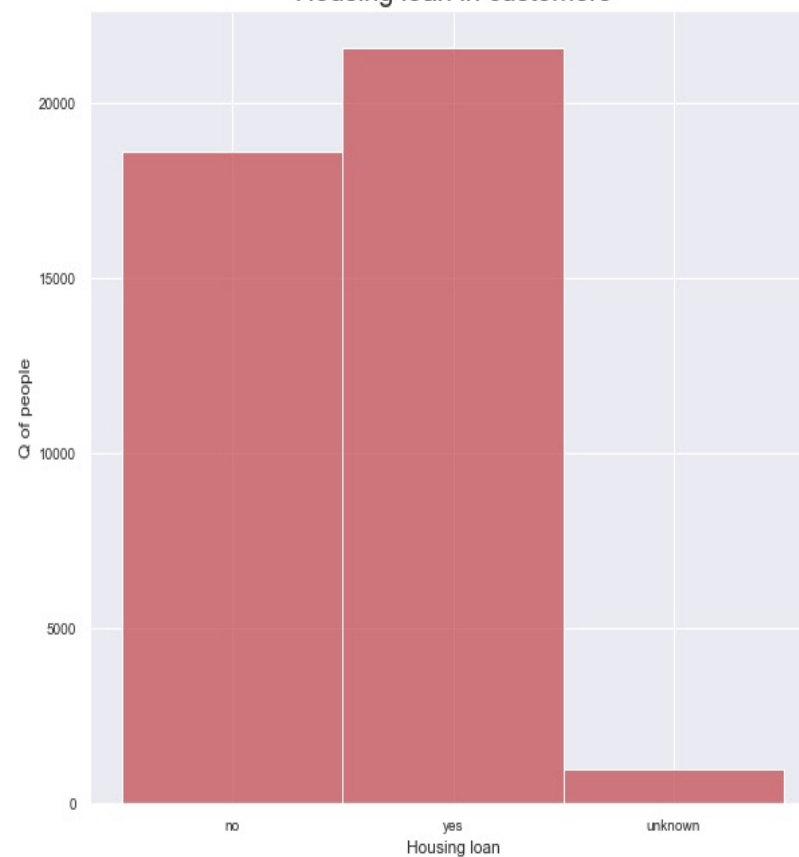
60% are married.  
 64% are admin, blue-collar or technician.  
 65% are 31-50 yo.  
 65% are graduated from HS, University or professional course.

## Customers and Default, Housing Loan and Personal Loan:

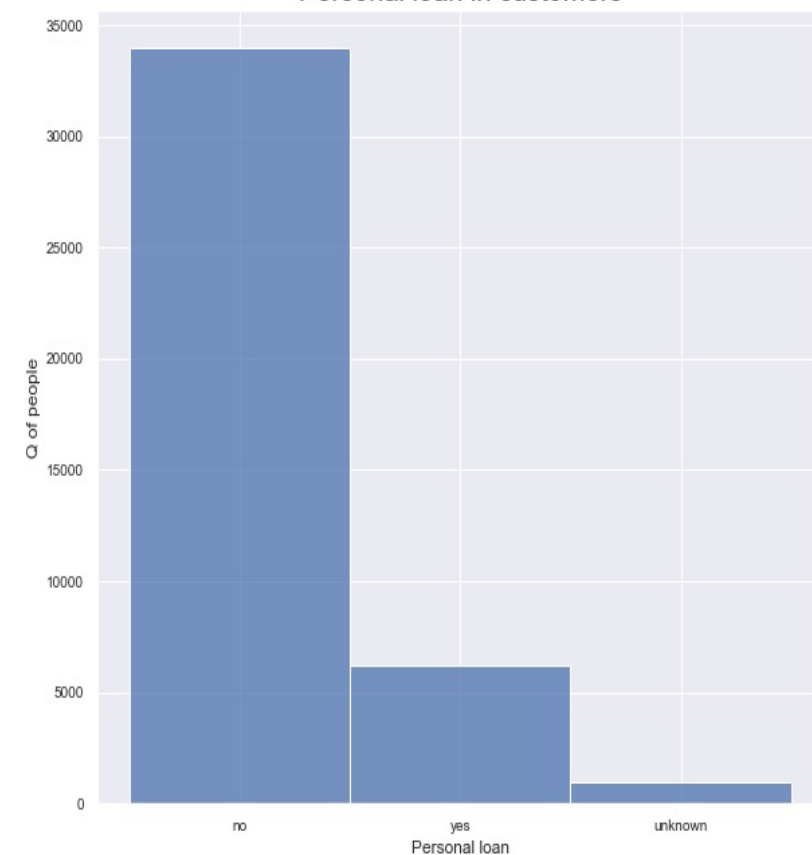
Default in customers

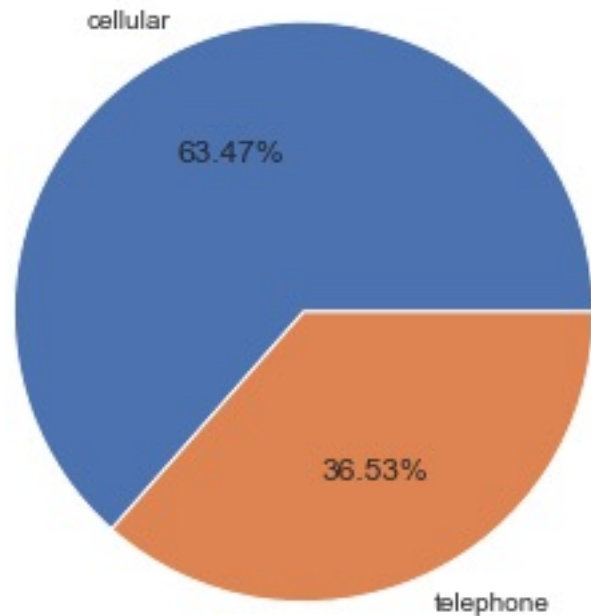


Housing loan in customers



Personal loan in customers





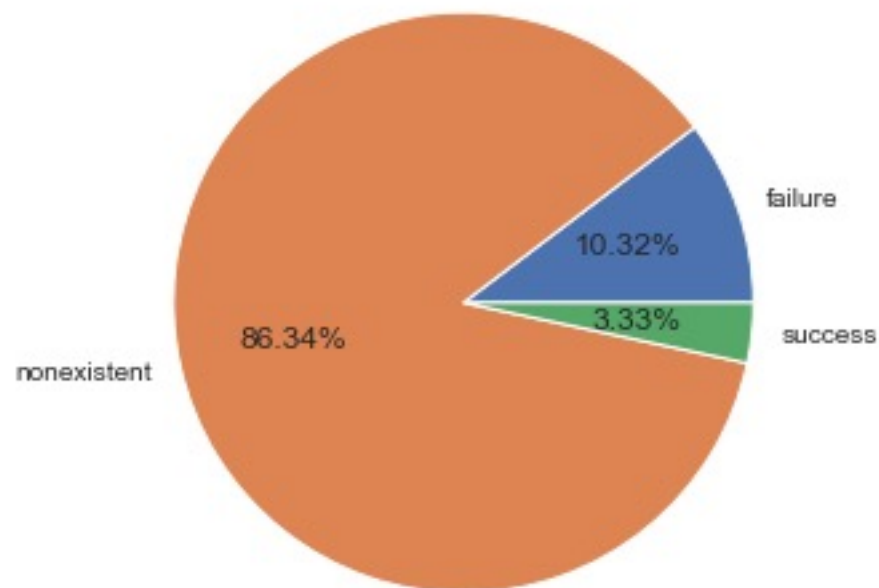
contact	
month	
may	33.429640
jul	17.417694
aug	14.999514
jun	12.911528
nov	9.956784

duration_range	
0-200	55.258813
201-400	27.347771
401-600	8.983199
601-800	4.144411
801-1000	1.944741
over 1000	2.321064

63% of contacts to customers are done by celular.

78% of contacts to customers are done in 4 months (may, july, august and june).

82% of calls last less 400 seconds.



```
campaign_range
0-3          81.463047
4-6          12.695445
7-9           3.185394
10-12         1.279499
More than 12  1.376615
Name: campaign, dtype: float64
```

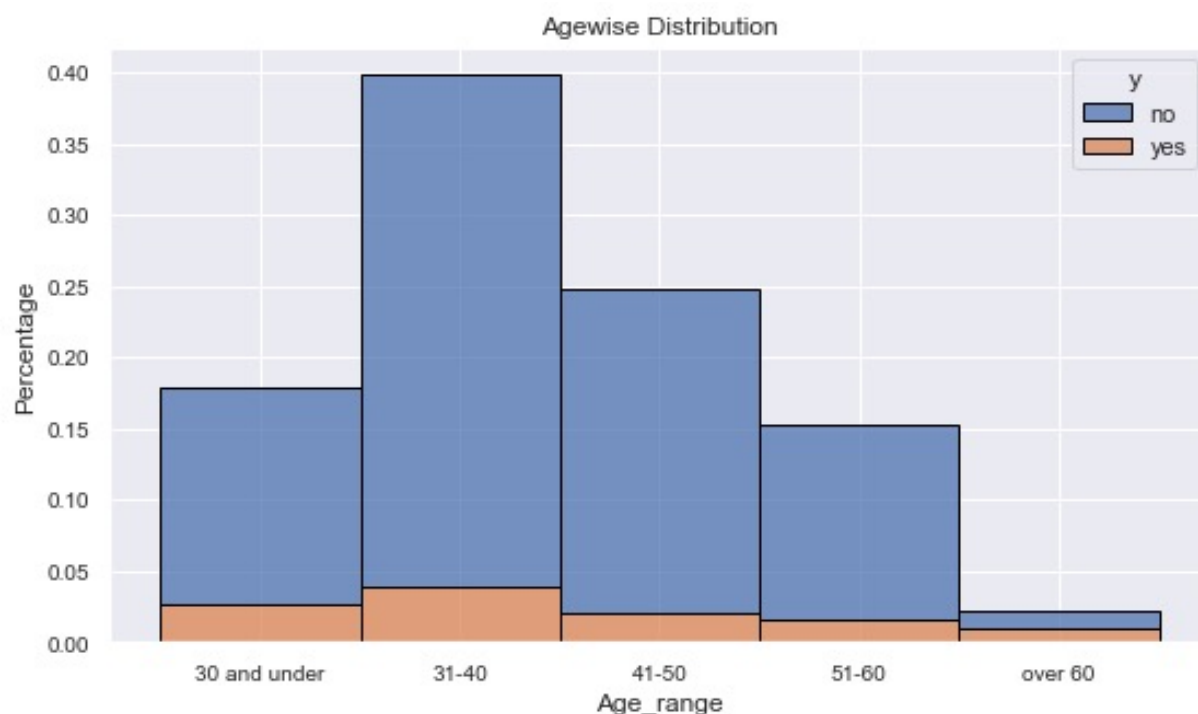
```
previous
0      86.343110
1      11.073614
2         1.830630
3         0.524425
4         0.169952
5         0.043702
6         0.012139
7         0.002428
..         .      ..
```

86% of customers don't have previous relation with the bank.

3,3% success of the 14% of customers who have previous contacts with the bank.



# Bivariial for Some features Age and job



People under 30 and over 60 yo are customers with higher acceptance rate. 15% and 45% respectively.

Accordinly, student and retired are top Jobs in acceptance rate.

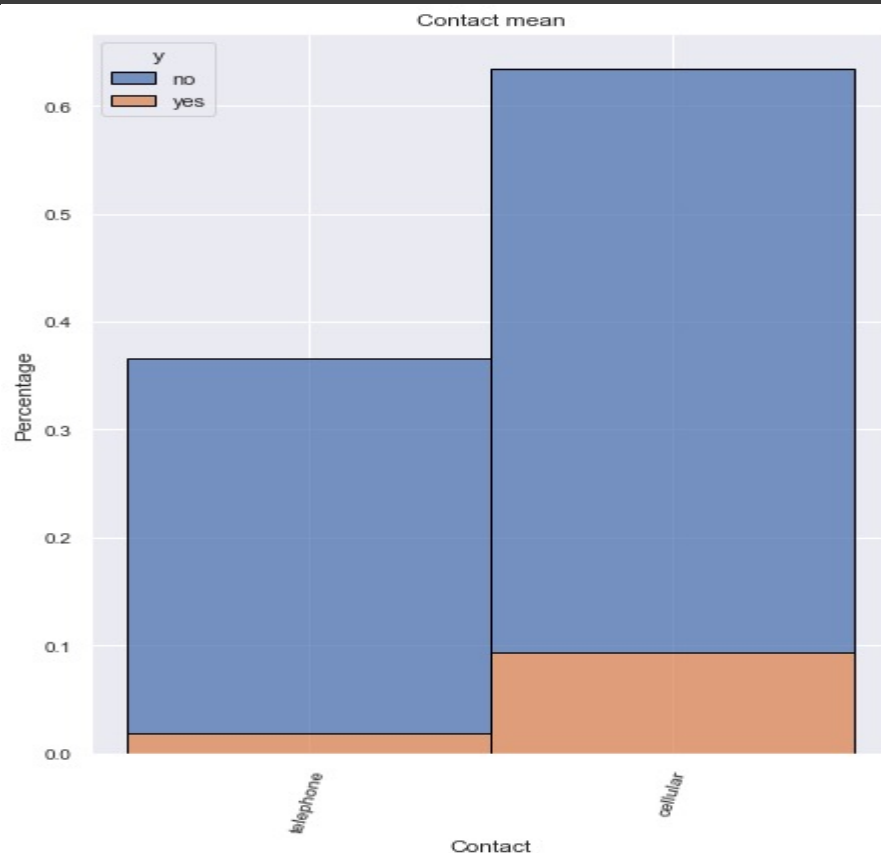
y	no	yes	totals	% no	% yes
job					
admin.	9070	1352	10422	87.027442	12.972558
blue-collar	8616	638	9254	93.105684	6.894316
entrepreneur	1332	124	1456	91.483516	8.516484
housemaid	954	106	1060	90.000000	10.000000
management	2596	328	2924	88.782490	11.217510
retired	1286	434	1720	74.767442	25.232558
self-employed	1272	149	1421	89.514426	10.485574
services	3646	323	3969	91.861930	8.138070
student	600	275	875	68.571429	31.428571
technician	6013	730	6743	89.173958	10.826042
unemployed	870	144	1014	85.798817	14.201183
unknown	293	37	330	88.787879	11.212121
totals	36548	4640	41188	88.734583	11.265417





# Bivariar for Some features

## Contact medium, month and call duration



y	no	yes	totals	% no	% yes
month					
apr	2093	539	2632	79.521277	20.478723
aug	5523	655	6178	89.397863	10.602137
dec	93	89	182	51.098901	48.901099
jul	6525	649	7174	90.953443	9.046557
jun	4759	559	5318	89.488530	10.511470
mar	270	276	546	49.450549	50.549451
may	12883	886	13769	93.565255	6.434745
nov	3685	416	4101	89.856133	10.143867
oct	403	315	718	56.128134	43.871866
sep	314	256	570	55.087719	44.912281

y	no	yes	totals	% no	% yes
duration_range					
0-300	28466	1518	29984	94.937300	5.062700
301-600	6302	1438	7740	81.421189	18.578811
601-900	1259	924	2183	57.672927	42.327073
901-1200	329	439	768	42.838542	57.161458
1201-1500	115	192	307	37.459283	62.540717
1501-1800	36	69	105	34.285714	65.714286
1801-2100	22	34	56	39.285714	60.714286
2101-2400	6	8	14	42.857143	57.142857
More than 2400	13	18	31	41.935484	58.064516

15% of people contacted by celular accepted the term deposit, while just 5% of the ones contacted by telephone. December, march, october and september have few customers, but with very high conversion rate (over 40%). Longer call, higher conversion rate. Consider that a small quantity of people have calls for more than 900 seconds.



With the explorative work done till now, we can think:

- 1) Young and old people (under 30 and over 60) have the highest conversion rates per age range. They represent the 20% of customers. Of course, there are some things to take into account such as risk averse.
- 2) Contact by celular is way more effective than telephone (15% vs 5% conversion rate).
- 3) In Bivarial analysis, we highlighted interesting relations we found at first sight and exploration. In the notebook is the analysis with the rest of features to Deep in.

# Thank You