



\$60,000 • 1,065 teams

## Santander Product Recommendation

Merger and Entry Deadline

Wed 26 Oct 2016

Wed 21 Dec 2016 (19 days to go)

### Dashboard

[Home](#)[Data](#)[Make a submission](#)

### Information

[Description](#)[Evaluation](#)[Rules](#)[Prizes](#)[Timeline](#)

### Forum

### Kernels

[New Script](#)[New Notebook](#)

### Leaderboard

### My Submissions

### Public Leaderboard

1. Tom Van de Wiele
2. 关山难越, 谁悲失路之人
3. lb overfitting snail
4. yoniko
5. idle\_speculation
6. BreakfastPirate
7. Jack (Japan)
8. Jared Turkewitz
9. Alejo y Miro
10. Sameh Faidi

### Competition Details » Get the Data » Make a submission

#### Data Files

File Name	Available Formats
sample_submission.csv	<a href="#">.zip (2.37 mb)</a>
test_ver2.csv	<a href="#">.zip (12.39 mb)</a>
train_ver2.csv	<a href="#">.zip (214.26 mb)</a>

In this competition, you are provided with 1.5 years of customers behavior data from Santander bank to predict what new products customers will purchase. The data starts at 2015-01-28 and has monthly records of products a customer has, such as "credit card", "savings account", etc. You will predict what **additional** products a customer will get in the last month, 2016-06-28, in addition to what they already have at 2016-05-28. These products are the columns named: ind\_(xyz)\_ult1, which are the columns #25 - #48 in the training data. You will predict what a customer will buy **in addition to what they already had at 2016-05-28**.

The test and train sets are split by time, and public and private leaderboard sets are split randomly.

**Please note: This sample does not include any real Santander Spain customers, and thus it is not representative of Spain's customer base.**

## File descriptions

## 1,737 Kernels

Detailed  
Cleaning/Visualization  
(Python)  
101 Votes / 9 days ago / Python

When Less is More.!!  
36 Votes / 11 days ago / Python

Comprehensive  
exploration and  
visualization-1  
27 Votes / 2 days ago / Python

Simple Exploration  
Notebook V3.0  
24 Votes / 3 days ago / Python

Detailed  
Cleaning/Visualization  
114 Votes / 19 days ago / R

Reduce Size of Dataset to  
1 GB  
17 Votes / 7 days ago / Python

## Forum (143 topics)

Difficult diversifying  
models?  
50 minutes ago

Reduce Size of Dataset to  
1 GB  
1 hour ago

first  
2 hours ago

When Less is More.!!23  
4 hours ago

No real customers?  
5 hours ago

When Santander deploys  
these models ...  
5 hours ago

teams

players

entries

- **train.csv** - the training set
- **test.csv** - the test set
- **sample\_submission.csv** - a sample submission file in the correct format

## Data fields

Column Name	Description
fecha_dato	The table is partitioned for this column
ncodpers	Customer code
ind_empleado	Employee index: A active, B ex employed, F filial, N not employee, P pasive
pais_residencia	Customer's Country residence
sexo	Customer's sex
age	Age
fecha_alta	The date in which the customer became as the first holder of a contract in the bank
ind_nuevo	New customer Index. 1 if the customer registered in the last 6 months.
antiguedad	Customer seniority (in months)
indrel	1 (First/Primary), 99 (Primary customer during the month but not at the end of the month)
ult_fec_cli_1t	Last date as primary customer (if he isn't at the end of the month)
indrel_1mes	Customer type at the beginning of the month ,1 (First/Primary customer), 2 (co-owner ),P (Potential),3 (former primary), 4(former co-owner)
tiprel_1mes	Customer relation type at the beginning of the month, A (active), I (inactive), P (former customer),R (Potential)
indresi	Residence index (S (Yes) or N (No) if the residence country is the same than the bank country)
indext	Foreigner index (S (Yes) or N (No) if the customer's birth country is different than the bank country)
conyuemp	Spouse index. 1 if the customer is spouse of an employee

Column Name	Description
canal_entrada	channel used by the customer to join
indfall	Deceased index. N/S
tipodom	Addres type. 1, primary address
cod_prov	Province code (customer's address)
nomprov	Province name
ind_actividad_cliente	Activity index (1, active customer; 0, inactive customer)
renta	Gross income of the household
segmento	segmentation: 01 - VIP, 02 - Individuals 03 - college graduated
ind_ahor_fin_ult1	Saving Account
ind_aval_fin_ult1	Guarantees
ind_cco_fin_ult1	Current Accounts
ind_cder_fin_ult1	Derivada Account
ind_cno_fin_ult1	Payroll Account
ind_ctju_fin_ult1	Junior Account
ind_ctma_fin_ult1	Más particular Account
ind_ctop_fin_ult1	particular Account
ind_ctpp_fin_ult1	particular Plus Account
ind_deco_fin_ult1	Short-term deposits
ind_deme_fin_ult1	Medium-term deposits
ind_dela_fin_ult1	Long-term deposits
ind_ecue_fin_ult1	e-account
ind_fond_fin_ult1	Funds
ind_hip_fin_ult1	Mortgage
ind_plan_fin_ult1	Pensions
ind_pres_fin_ult1	Loans
ind_reca_fin_ult1	Taxes
ind_tjcr_fin_ult1	Credit Card
ind_valo_fin_ult1	Securities
ind_viv_fin_ult1	Home Account
ind_nomina_ult1	Payroll
ind_nom_pens_ult1	Pensions
ind_recibo_ult1	Direct Debit