What does Receipt Bank* do?

not a bank

Dataset (9K receipts)

```
In [1]:  import pandas as pd

In [2]:  receipts = pd.read_csv('receipts.csv', nrows=100)

In [8]:  receipts
```

	id	created_at	currency_code	total_amount	vat_amount	date	due_date	invoice_number	received_via	supplier_name	ocr
0	49927358	2017-01-25 14:01:42.251023	USD	585.70	0.00	2017-01-02	2017-01-30	386764	via_supplier_email	Intercom	clou
1	49897860	2017-01-25 09:11:29.719365	USD	1013.00	0.00	2017-01-09	NaN	NaN	via_supplier_email	Mixpanel	clou
2	49845841	2017-01-24 21:14:12.473547	GBP	618.47	0.00	2017-01-24	NaN	NaN	via_mobile	Brasserie Blanc	clou
3	49825653	2017-01-24 16:31:36.352851	GBP	2400.00	400.00	2017-01-24	2017-01-31	03626	via_supplier_email	Finatal	clou
4	49825652	2017-01-24 16:31:36.121024	GBP	6000.00	1000.00	2017-01-24	2017-01-31	03628	via_supplier_email	Finatal	clou
5	49820985	2017-01-24 15:47:10.967728	GBP	40.35	0.00	2017-01-24	NaN	NaN	via_mobile	Sainsbury's	clou

Images

Tom's Kitchen

27 Cale Street
London
SW3 3QP
Tel. +44 207 349 0202
www.tomskitchen.co.uk

Server:Letuppe, Emilie
12:48pm 11/09/12
Table No:31   Covers: 5

1 Belu Sparkling	£3.50
1 Slimline Tonic	£2.00
1 l Orange Juice	£4.25
2 Diet Coke	£6.00
1 Calves Liver	£18.50
1 Pie of Day	£12.50
1 Chick Schnitzel	£19.00
1 Pork Belly	£18.00
1 Monkfish	£18.50
2 Chips	£8.00
1 Spinach	£4.00
1 Rocket + Parmesan	£4.00
2 Lemon Tart	£14.00
1 Chocolate Tart	£7.00
2 Vanilla Cheesecake	£14.00

Opt Svc Ch:	£19.16
Amt.Due	**£172.41**

VAT Included	£25.53

Includes VAT at current rate.
Please note our Weekend Breakfast
Menu will finish @ 12.30pm and
our Weekend Lunch Menu begins.
VAT No:927 47 6685 Inv No:33169

*** CUSTOMER COPY ***

TOM@S KITCHEN
27 CALE STREET
LONDON
SW3 3QP
THANK YOU

14:18:23   11/09/12
MID: ********92747   TID: ****0817

Visa
***********8960
01
A0000000031010
SALE   CHIP
                        £172.41

PIN VERIFIED
APPROVED
AUTH CODE: 000587
44C780A954DBA058   TC
S:2   DIAG: 73

Please debit my account
with the total amount
PLEASE RETAIN RECEIPT

The WestRoom
3 Melville Place
Edinburgh

Till 1
Goose   30 Oct 2012 17:37

Table: 1 c: 2   Acc No: 8288

3   Speciality Tea   =   6.90
_____

Total   £6.90
_____

Service charge is not included.
A 10% SERVICE CHARGE IS INCLUDED
ON PARTIES OF 6 OR MORE

PDF

- Receipt Bank

  settingsbillingcollaborators
  - User Settings
  - Sign Out
- General
- Plans & Billing
- Collaborators
- Invoices
- Tell a Friend

- Support
- Twitter
- Blog
- Docs
- Shortcuts

© 2015 Bugsnag Inc.

Print invoice

Bugsnag Invoice

Bugsnag Inc., 30 7th St., San Francisco, CA, 94103, USA

Summary

Invoice Number	Receipt Bank 2015-02-05 - 2015-03-05
Bill to	Receipt Bank
Statement Date	Feb 5, 2015
Billing Period	Feb 5, 2015 to Mar 4, 2015
Billing Plan	Growth Plan

Charges

Name on Card	J D Ryall
Card Number	**** **** **** 8112
Status	Paid
Plan Base Price	$99.00 USD
Total	$99.00 USD
Charged to Card	$99.00 USD

- Terms of Service
- Pricing
- Enterprise

RECEIPT / TAX INVOICE

capsule

To
Receipt Bank
United Kingdom

Zestia Limited
The Old Workshop
12b Kennerleys Lane
Wilmslow
Cheshire SK9 5EQ
United Kingdom

Email
info@zestia.com

VAT Number
GB 934502242

Company Number
6418281

Invoice Date
Jul 17, 2012

Invoice Number
48549-00001

Account URL
https://receiptbankdaniel.capsulecrm.com

Description	Amount GBP
Capsule subscription for Jul 17, 2012 to Aug 16, 2012. 1 users @ £8	8.00
Less upgrade offer discount	-4.00
Subtotal	4.00
VAT	0.80
Total GBP	**4.80**

Automatic debit - no action required
No action is required. Payment for this invoice has been taken from your card ************3826.

Text

```json
[
    {
        "textAnnotations": [
            {
                "locale": "en",
                "description": "ma\nTOTAL VARIABLE PRINTING AND MAILING SOLUTIONS\nBerwick Hill Dairy Berwick St Leonard Salisbury Wiltshire SP3 5SN\nsales@mailandprint.co.uk Tel[...skipped…]hurchill Plaza, Churchill Way, Basingstoke,\nHampshire RG21 7GL. To whom payment must be made and whose receipt therefore alone is valid.\n",
                "boundingPoly": {
                    "vertices": [
                        {
                            "x": 87,
                            "y": 186
                        },
                        {
                            "x": 2325,
                            "y": 186
                        },
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE pdf2xml SYSTEM "pdf2xml.dtd">

<pdf2xml producer="poppler" version="0.51.0">
<page number="1" position="absolute" top="0" left="0" height="1263" width="892">
    <fontspec id="0" size="16" family="Times" color="#000000"/>
<text top="76" left="544" width="11" height="23" font="0">&gt;</text>
<text top="100" left="52" width="292" height="23" font="0">Date: Tue, Sep 11, 2012 at 2:53 PM</text>
<text top="123" left="52" width="398" height="23" font="0">Subject: Your Booking Confirmation
9NWRJ8FG</text>
<text top="146" left="52" width="32" height="23" font="0">To:<a href="mailto:MICHAEL.WOOD@gmail.com">
</a></text>
<text top="146" left="83" width="249" height="23" font="2"><a
href="mailto:MICHAEL.WOOD@gmail.com">MICHAEL.WOOD@gmail.com</a></text>
<text top="169" left="52" width="33" height="23" font="0">Cc:<a href="mailto:MICHAEL.WOOD@gmail.com">
</a></text>
<text top="169" left="84" width="249" height="23" font="2"><a
href="mailto:MICHAEL.WOOD@gmail.com">MICHAEL.WOOD@gmail.com</a></text>
<text top="393" left="55" width="472" height="19" font="3"><b>*This is confirmation of your booking and
is not your travel ticket.</b></text>
<text top="438" left="55" width="134" height="23" font="0">Dear Mr WOOD,</text>
```

Problems

Currency (easy)

Recognize the currency

(Мотика: There are many dollars around the words - USD, AUD, ZAR)

TOTAL (medium)

What's the total (as a number).

Supplier (hard)

Who issued the receipt.

We don't need 100% accurate guess, e.g. "Starbucks", "Starbucks Ltd", "Starbucks Limited Company" are all correct guesses.

Supplier database in dataset.

MAIL BOXES ETC
LONDON SW7 3SS
TEL 0207 5812825
VAT 656 6507 11

REG 08-10-2012(MON) 15:19
JOHN   MC#01   098735

2 STAT   ·0.90
  TA1   ·0.75
  VAT   ·0.15
  TL   ·0.90
  CASH   ·10.00
  CG   ·9.10

Good luck!