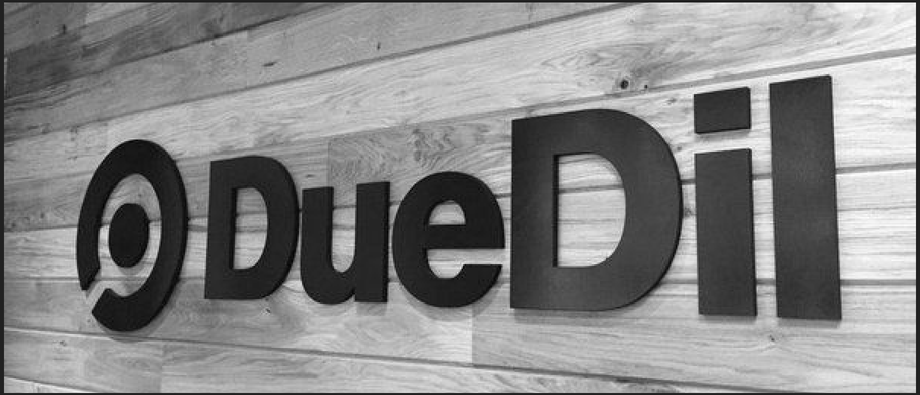
10 things I learned about writing data pipelines in Python and Spark

#7 will blow your mind!

In the next 30 minutes

- Patterns that you can apply today
- There will be <code>
- No machine learning - yet
- Use with caution*
- Mostly Python
- Who is this talk for?

Background

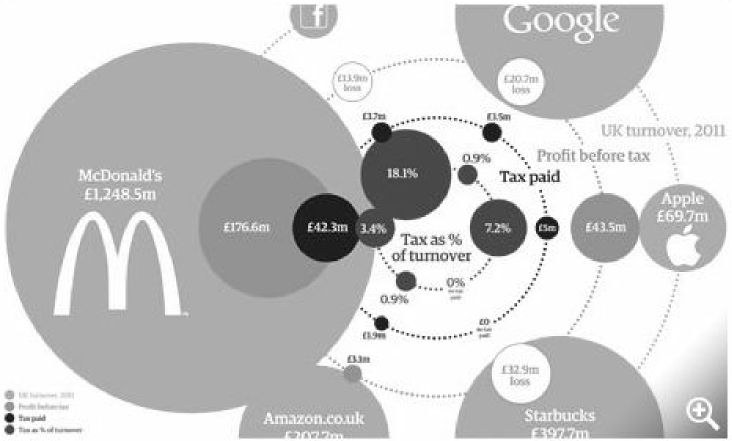


How much tax do Starbucks, Facebook and the biggest US companies pay in the UK

How much do the biggest US companies in the UK pay in tax?

• Download the data

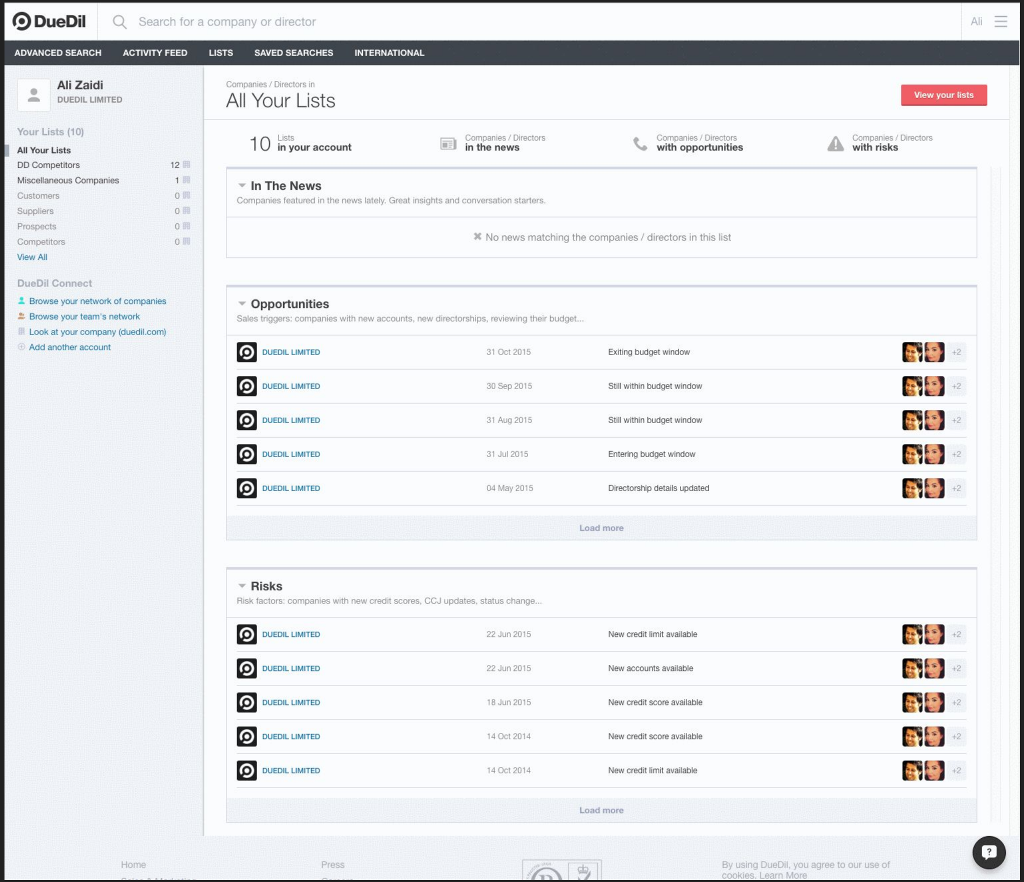
• More data journalism and data visualisations from the Guardian



How much tax is paid by major US companies in the UK? Click image to embiggen

Background

Enhanced Financials


DueDil — Search for a company or director — Ali

ADVANCED SEARCH | ACTIVITY FEED | LISTS | SAVED SEARCHES | INTERNATIONAL

Ali Zaidi
DUEDIL LIMITED

Your Lists (10)

All Your Lists	
DD Competitors	12
Miscellaneous Companies	1
Customers	0
Suppliers	0
Prospects	0
Competitors	0

View All

DueDil Connect
- Browse your network of companies
- Browse your team's network
- Look at your company (duedil.com)
- Add another account

Companies / Directors in
All Your Lists

View your lists

10 Lists **in your account** | Companies / Directors **in the news** | Companies / Directors **with opportunities** | Companies / Directors **with risks**

In The News
Companies featured in the news lately. Great insights and conversation starters.

✖ No news matching the companies / directors in this list

Opportunities
Sales triggers: companies with new accounts, new directorships, reviewing their budget...

DUEDIL LIMITED	31 Oct 2015	Exiting budget window	+2
DUEDIL LIMITED	30 Sep 2015	Still within budget window	+2
DUEDIL LIMITED	31 Aug 2015	Still within budget window	+2
DUEDIL LIMITED	31 Jul 2015	Entering budget window	+2
DUEDIL LIMITED	04 May 2015	Directorship details updated	+2

Load more

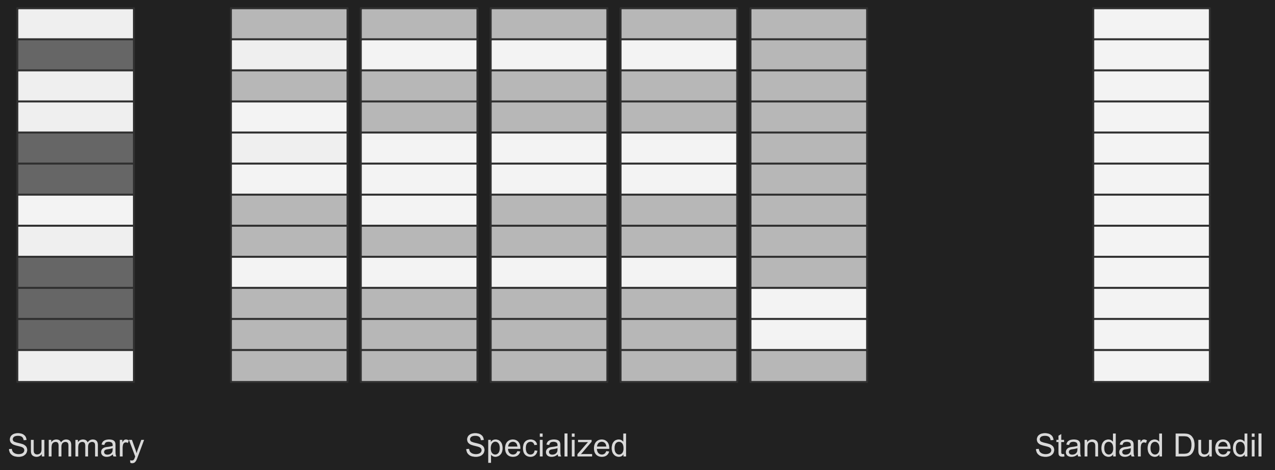
Risks
Risk factors: companies with new credit scores, CCJ updates, status change...

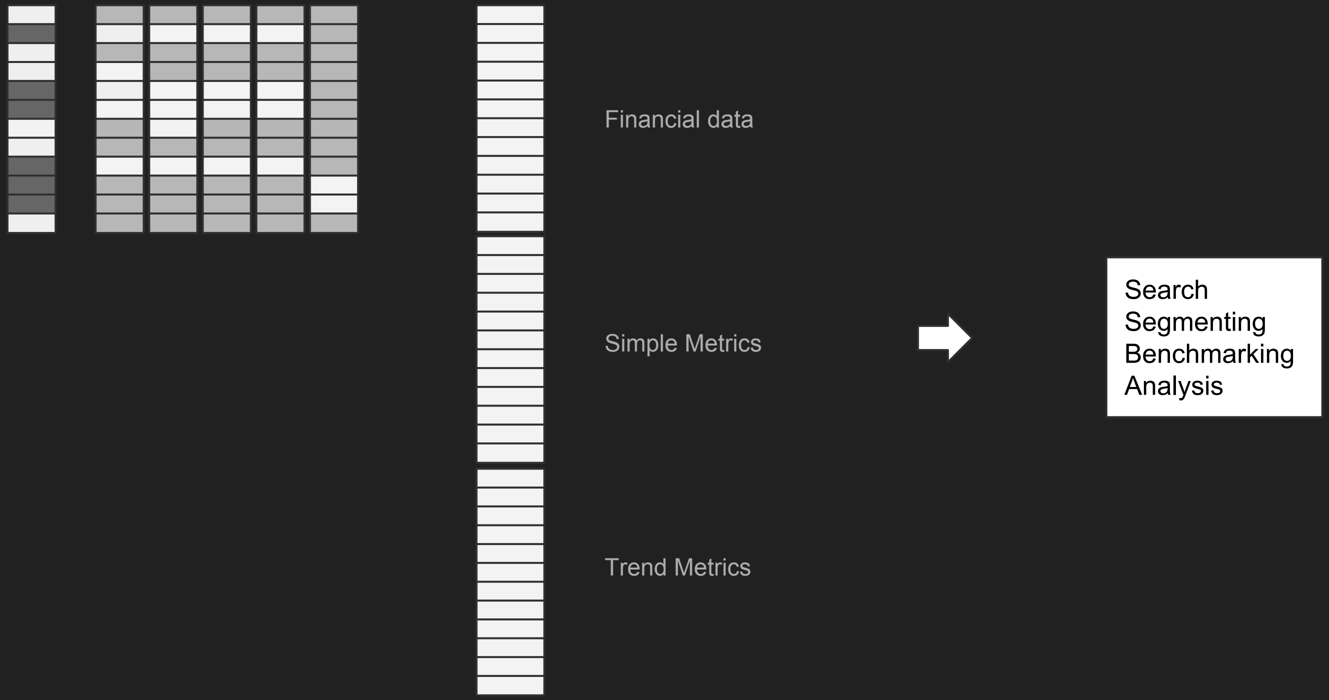
DUEDIL LIMITED	22 Jun 2015	New credit limit available	+2
DUEDIL LIMITED	22 Jun 2015	New accounts available	+2
DUEDIL LIMITED	18 Jun 2015	New credit score available	+2
DUEDIL LIMITED	14 Oct 2014	New credit score available	+2
DUEDIL LIMITED	14 Oct 2014	New credit limit available	+2

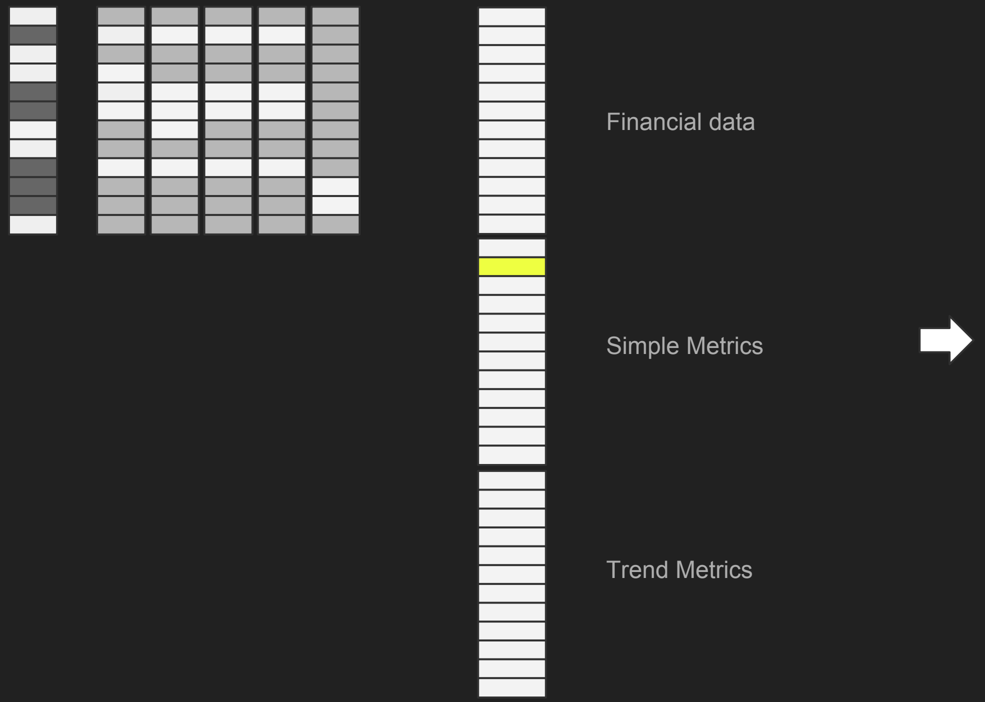
Load more

Home | Press | By using DueDil, you agree to our use of cookies. Learn More

The Data




Financial data
Simple Metrics
Trend Metrics
Search
Segmenting
Benchmarking
Analysis



Search
Segmenting
Benchmarking
Analysis

```python
def gearing(liabilities, shareholder_funds):
    return (liabilities / shareholder_funds) * 100
```

```python
def gearing(rec):
    return (rec['liabilities'] /
            rec['shareholder_funds']) * 100
```

```
def gearing(rec):
    return (rec['liabilities'] * 1.0 /
            rec['shareholder_funds']) * 100
```

```
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    return (rec['liabilities'] * 1.0 /
            rec['shareholder_funds']) * 100
```

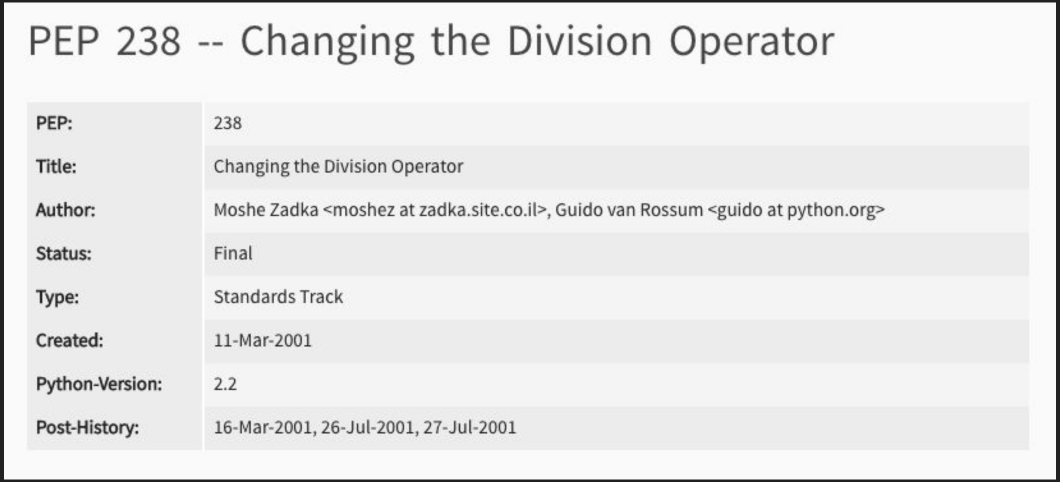
```python
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['shareholder_funds'] == 0:
        return None
    return (rec['liabilities'] * 1.0 /
            rec['shareholder_funds']) * 100
```

```python
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['shareholder_funds'] == 0:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] * 1.0 /
            rec['shareholder_funds']) * 100
```

#1 Cleaning code will eat compute code and code quality

```
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['shareholder_funds'] == 0:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] * 1.0 /
            rec['shareholder_funds']) * 100
```

```python
from __future__ import division

def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['shareholder_funds'] == 0:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] /
            rec['shareholder_funds']) * 100
```



PEP 238 -- Changing the Division Operator

PEP:	238
Title:	Changing the Division Operator
Author:	Moshe Zadka <moshez at zadka.site.co.il>, Guido van Rossum <guido at python.org>
Status:	Final
Type:	Standards Track
Created:	11-Mar-2001
Python-Version:	2.2
Post-History:	16-Mar-2001, 26-Jul-2001, 27-Jul-2001

#2 Disable Python’s default Integer division*

```
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['shareholder_funds'] == 0:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] /
            rec['shareholder_funds']) * 100
```

```
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    try:
        return (rec['liabilities'] /
                rec['shareholder_funds']) * 100
    except ZeroDivisionError:
        return None
```

```
@handle_div_by_0
def gearing(rec):
    if 'shareholder_funds' not in rec or
       'liabilities' not in rec:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] /
            rec['shareholder_funds']) * 100
```

```python
from functools import wraps

def handle_div_by_0(f):
    @wraps(f)
    def wrapper(*arg, **kwargs):
        try:
            return f(*arg, **kwargs)
        except ZeroDivisionError:
            return None
    return wrapper
```

#3 Handle ZeroDivisionError in decorators**

```
def gearing(rec):
    if 'shareholder_funds' not in  rec or
       'liabilities' not in  rec:
        return None
    if rec['liabilities'] is None or
       rec['shareholder_funds'] is None:
        return None
    return (rec['liabilities'] /
            rec['shareholder_funds']) * 100
```

```python
def gearing(shareholder_funds, liabilities, **kwargs):
    if liabilities is None or shareholder_funds is None:
        return None
    return (liabilities / shareholder_funds) * 100
```

```python
def gearing(shareholder_funds, liabilities, **kwargs):
    '''Replace gearing(rec) with gearing(**rec)'''
    if liabilities is None or shareholder_funds is None:
        return None
    return (liabilities / shareholder_funds) * 100
```

```python
def gearing(shareholder_funds, liabilities, **kwargs):
    '''Replace gearing(rec) with gearing(**rec)
    And this would still not work :P
    '''
    if liabilities is None or shareholder_funds is None:
        return None
    return (liabilities / shareholder_funds) * 100
```

```
def gearing(shareholder_funds=None, liabilities=None, **kwargs):
    if liabilities is None or shareholder_funds is None:
        return None
    return (liabilities / shareholder_funds) * 100
```

#4 Use **kwargs to unpack records into local variables

```
def gearing(shareholder_funds=None, liabilities=None, **kwargs):
    if liabilities is None or shareholder_funds is None :
        return None
    return (liabilities / shareholder_funds) * 100
```

```python
def gearing(shareholder_funds=None, liabilities=None, **kwargs):
    if None in [shareholder_funds, liabilities] :
        return None
    return (liabilities / shareholder_funds) * 100
```

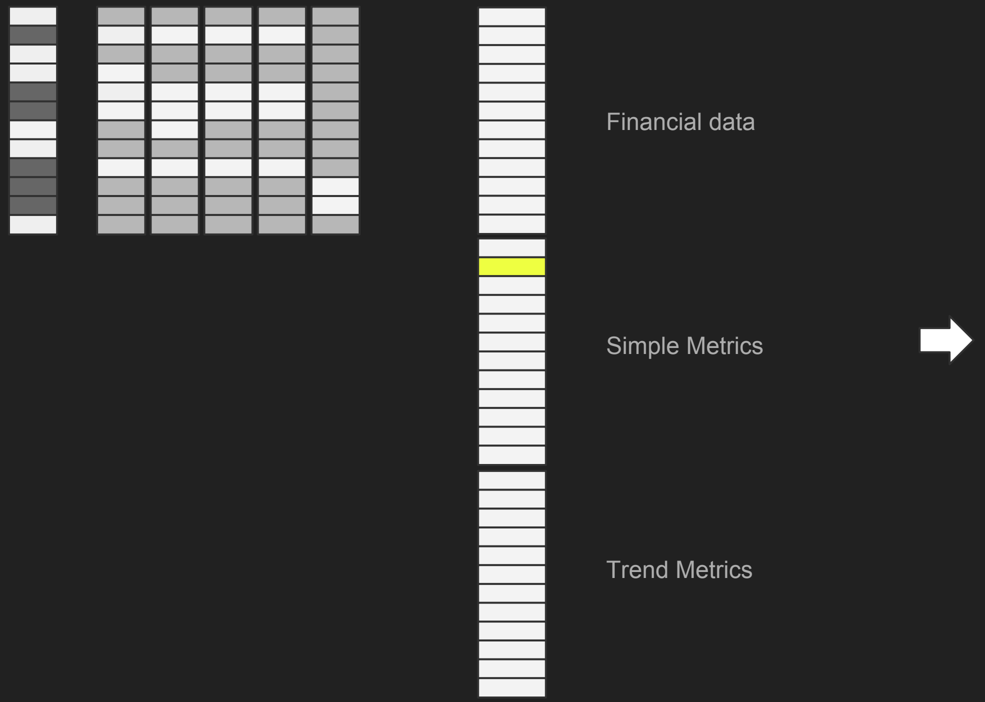
#5 Refactor checks for many variables having a value to “in”

```python
def gearing(shareholder_funds=None, liabilities=None, **kwargs):
    if None in [shareholder_funds, liabilities]:
        return None
    return (liabilities / shareholder_funds) * 100
```

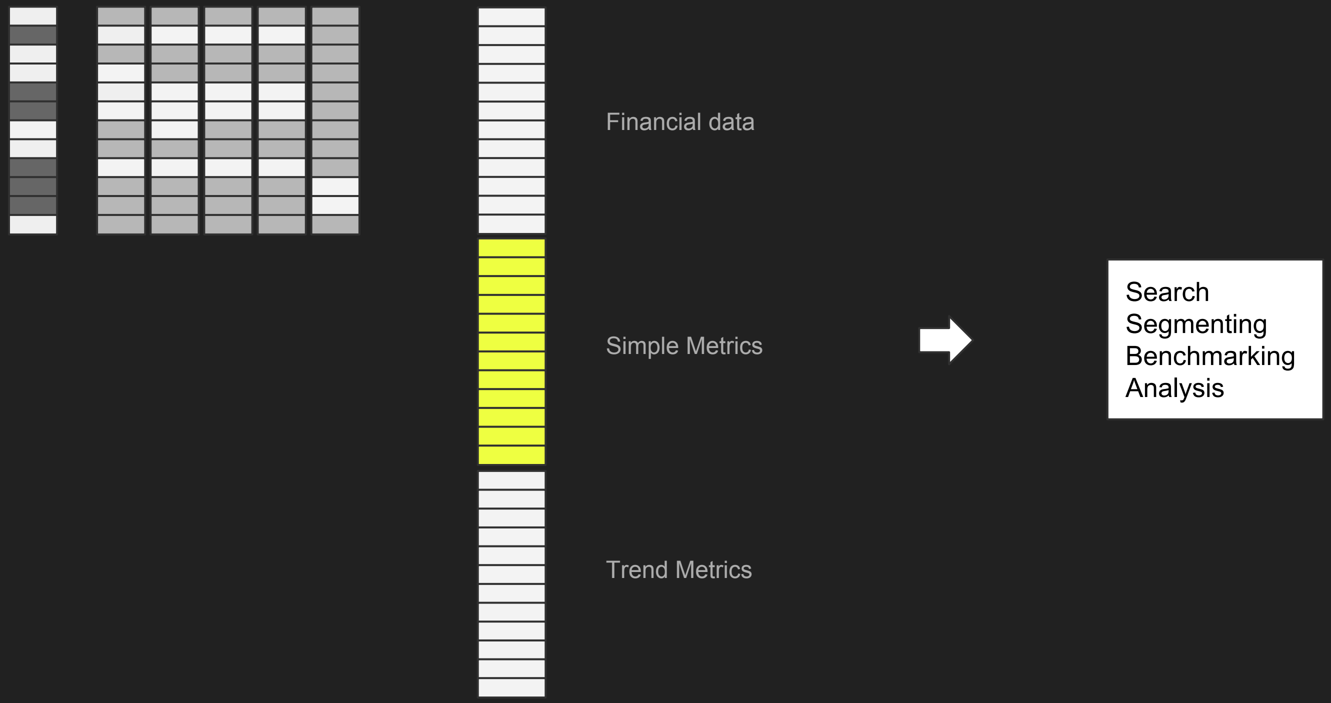
```
@handle_none_args
def gearing(shareholder_funds, liabilities, **kwargs):
    return (liabilities / shareholder_funds) * 100
```

```python
from functools import wraps
from inspect import getargspec

def handle_none_args(f):
    @wraps(f)
    def wrapper(**kwargs):
        for arg in getargspec(f)[0]:
            if kwargs.get(arg) is None:
                return None
        return f(**kwargs)
    return wrapper
```

#6 Be a craftsman but don’t be a perfectionist



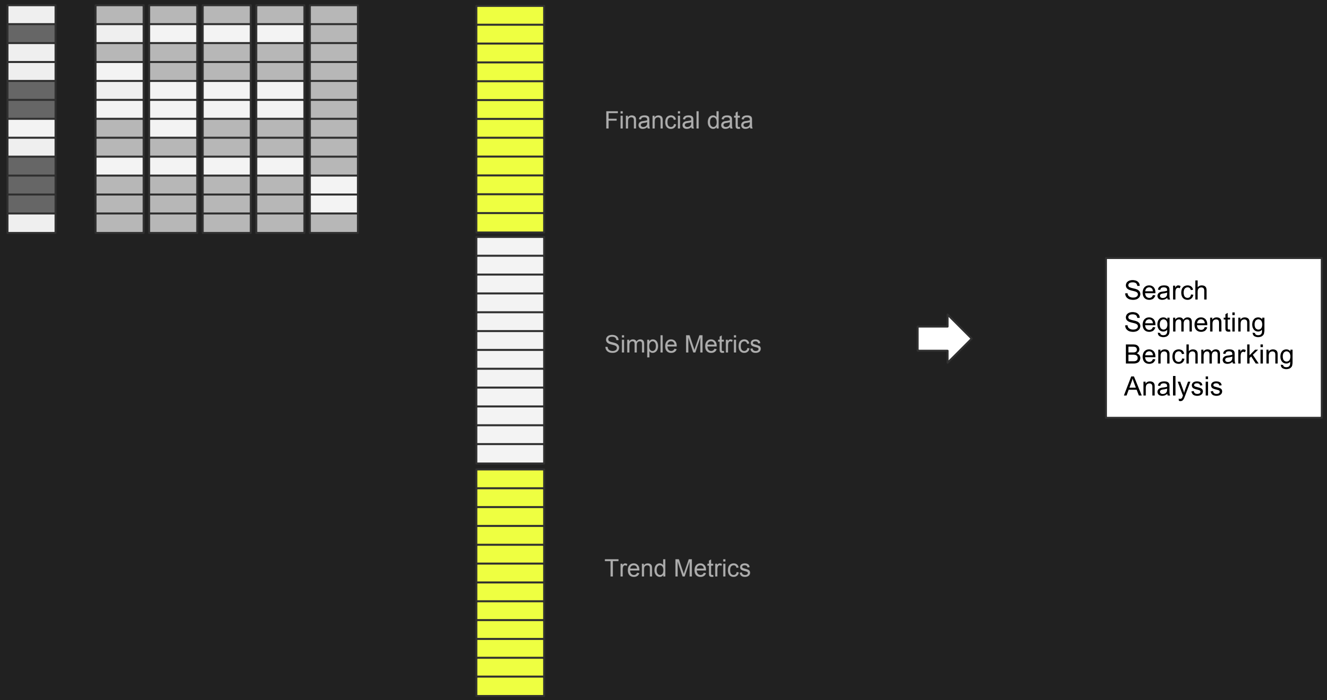
Search
Segmenting
Benchmarking
Analysis


Financial data
Simple Metrics
Trend Metrics
Search
Segmenting
Benchmarking
Analysis

```python
def calc_metrics(rec):
    '''Without kwarg unpacking'''
    rec['gearing'] = gearing(
        rec['shareholder_funds'],
        rec['liabilities']
    )
    # 200 more...
```

```
METRICS_MAP = [
    ('gearing', gearing),
    # 100 more
]

def calc_metrics(rec):
    for field, func in METRICS_MAP:
        rec[field] = func(**rec)
```

#7 Mapping dictionaries are awesome


Financial data
Simple Metrics
Trend Metrics
Search
Segmenting
Benchmarking
Analysis

```
TREND_METRICS_MAP = {
    ...
    prev_3: {
        'turnover': (cagr3, 'turnover'),
        # 50 more
    }
    ...
}
```

```
def sliding_window(iterable,  size=2):
    iters = tee(iterable, size)
    for i in xrange(1, size):
        for each in iters[i:]:
            next(each, None)
    return izip(*iters)
```

```
def sliding_window(iterable,  size=2):
    iters = tee(iterable, size)
    for i in xrange(1, size):
        for each in iters[i:]:
            next(each, None)
    return izip(*iters)

sliding_window(accounts)
sliding_window(accounts, 3)
```

```
def sliding_window(iterable, size):
    ...

prev_2 = partial(sliding_window, size=2)
prev_3 = partial(sliding_window, size=3)

prev_2(accounts)
prev_3(accounts)
```

#8 Use partials to make APIs more accessible

```
>>> fns = []
>>> for i in range(3):
...     fns.append(lambda: i)
...
>>> for fn in fns:
>>>     print fn()
>>> fns = []
>>> for i in range(3):
...     fns.append(lambda: i)
...
>>> for fn in fns:
>>>     print fn()
2
2
2
```

```
>>> fns = []
>>> for i in range(3):
...   def f(i=i):
...     return i
...   fns.append(f)
...
>>> for fn in fns:
...   print fn()
0
1
2
```

#9 Be aware of common Python gotchas

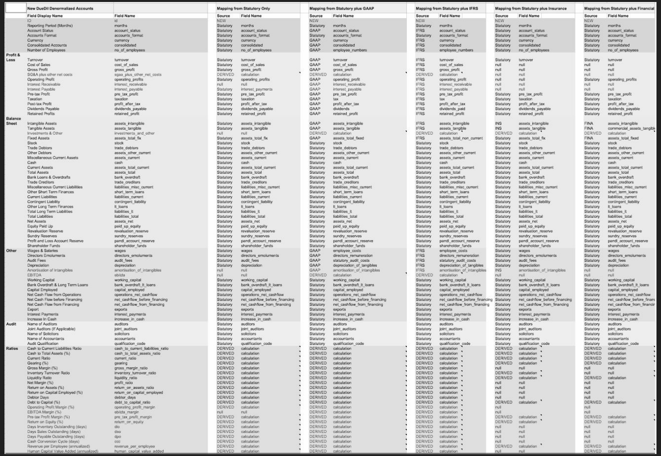
```
>>> type(1)
int
>>> type(1.0)
float
```

```
>>> data = [{'x': 353463435324.0}, {'x': 4534123897.0}]

>>> rdd = sc.parallelize(data)

>>> fields = [types.StructField('a', types. FloatType())]

>>> schema = types.StructType(fields)

>>> sql.jsonRDD(rdd, schema=schema).collect()
>>> data = [{'x': 353463435324.0}, {'x': 4534123897.0}]

>>> rdd = sc.parallelize(data)

>>> fields = [types.StructField('a', types.FloatType())]

>>> schema = types.StructType(fields)

>>> sql.jsonRDD(rdd, schema=schema).collect()
[Row(a=353463435264.0), Row(a=4534124032.0)]
```

#10 Be aware of common coding gotchas

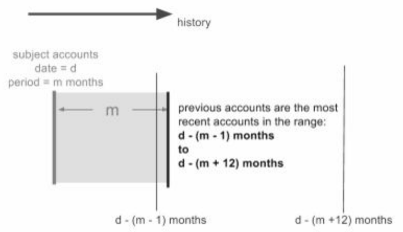
WhatsApp ups its group chat limit to 256 users

by MIX — 3 months ago in APPS



	New DueDil Denormalized Accounts		Mapping from Statutory Only		Mapping from Statutory plus GAAP		Mapping from Statutory plus IFRS		Mapping from Statutory plus Insurance		Mapping from Statutory plus Financial	
	Field Display Name	Field Name	Source	Field Name	Source	Field Name	Source	Field Name	Source	Field Name	Source	Field Name
	ID	id	NEW		NEW		NEW		NEW		NEW	
	Reporting Period (Months)	months	Statutory	months	Statutory	months	Statutory	months	Statutory	months	Statutory	months
	Account Status	account_status	Statutory	account_status	GAAP	account_status	IFRS	account_status	Statutory	account_status	Statutory	account_status
	Accounts Format	accounts_format	Statutory	accounts_format	Statutory	accounts_format	Statutory	accounts_format	Statutory	accounts_format	Statutory	accounts_format
	Currency	currency	Statutory	currency	GAAP	currency	IFRS	currency	Statutory	currency	Statutory	currency
	Consolidated Accounts	consolidated	Statutory	consolidated	GAAP	consolidated	IFRS	consolidated	Statutory	consolidated	Statutory	consolidated
	Number of Employees	no_of_employees	Statutory	no_of_employees	GAAP	employee_numbers	IFRS	employee_numbers	Statutory	no_of_employees	Statutory	no_of_employees
Profit & Loss	Turnover	turnover	Statutory	turnover	GAAP	turnover	IFRS	turnover	Statutory	turnover	Statutory	turnover
	Cost of Sales	cost_of_sales	Statutory	cost_of_sales	GAAP	cost_of_sales	IFRS	cost_of_sales	null	null	null	null
	Gross Profit	gross_profit	Statutory	gross_profit	GAAP	gross_profit	IFRS	gross_profit	null	null	null	null
	SG&A plus other net costs	sgaa_plus_other_net_costs	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Operating Profit	operating_profits	Statutory	operating_profits	GAAP	operating_profit	IFRS	operating_profit	null	null	Statutory	operating_profits
	Interest Receivable	interest_receivable	null	null	GAAP	interest_receivable	IFRS	interest_receivable	null	null	null	null
	Interest Payable	interest_payable	Statutory	interest_payments	GAAP	interest_payable	IFRS	interest_payable	null	null	null	null
	Pre-tax Profit	pre_tax_profit	Statutory	pre_tax_profit	GAAP	pre_tax_profit	IFRS	pre_tax_profit	Statutory	pre_tax_profit	Statutory	pre_tax_profit
	Taxation	taxation	Statutory	taxation	GAAP	tax	IFRS	tax	Statutory	taxation	Statutory	taxation
	Post-tax Profit	profit_after_tax	Statutory	profit_after_tax	GAAP	profit_after_tax	IFRS	profit_after_tax	Statutory	profit_after_tax	Statutory	profit_after_tax
	Dividends Payable	dividends_payable	Statutory	dividends_payable	GAAP	dividends	IFRS	dividends_paid	Statutory	dividends_payable	Statutory	dividends_payable
	Retained Profits	retained_profit	Statutory	retained_profit	GAAP	retained_profit	IFRS	retained_profit	Statutory	retained_profit	Statutory	retained_profit
Balance Sheet	Intangible Assets	assets_intangible	Statutory	assets_intangible	GAAP	assets_intangible	IFRS	assets_intangible	INS	assets_intangible	FINA	assets_intangible
	Tangible Assets	assets_tangible	Statutory	assets_tangible	GAAP	assets_tangible	IFRS	assets_tangible	INS	assets_tangible	FINA	commercial_assets_tangibl
	Investments & Other	investments_and_other	null	null	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Fixed Assets	assets_total_fix	Statutory	assets_total_fix	GAAP	assets_total_fixed	IFRS	assets_total_non_current	Statutory	assets_total_fix	FINA	assets_total_fixed
	Stock	stock	Statutory	stock	Statutory	stock	Statutory	stock	Statutory	stock	Statutory	stock
	Trade Debtors	trade_debtors	Statutory	trade_debtors	Statutory	trade_debtors	Statutory	trade_debtors	Statutory	trade_debtors	Statutory	trade_debtors
	Other Debtors	assets_other_current	Statutory	assets_other_current	Statutory	assets_other_current	Statutory	assets_other_current	Statutory	assets_other_current	Statutory	assets_other_current
	Miscellaneous Current Assets	assets_current	Statutory	assets_current	Statutory	assets_current	Statutory	assets_current	Statutory	assets_current	Statutory	assets_current
	Cash	cash	Statutory	cash	Statutory	cash	Statutory	cash	Statutory	cash	Statutory	cash
	Current Assets	assets_total_current	Statutory	assets_total_current	Statutory	assets_total_current	Statutory	assets_total_current	Statutory	assets_total_current	Statutory	assets_total_current
	Total Assets	assets_total	Statutory	assets_total	Statutory	assets_total	Statutory	assets_total	Statutory	assets_total	Statutory	assets_total
	Bank Loans & Overdrafts	bank_overdraft	Statutory	bank_overdraft	Statutory	bank_overdraft	Statutory	bank_overdraft	Statutory	bank_overdraft	Statutory	bank_overdraft
	Trade Creditors	trade_creditors	Statutory	trade_creditors	Statutory	trade_creditors	Statutory	trade_creditors	Statutory	trade_creditors	Statutory	trade_creditors
	Miscellaneous Current Liabilities	liabilities_misc_current	Statutory	liabilities_misc_current	Statutory	liabilities_misc_current	Statutory	liabilities_misc_current	Statutory	liabilities_misc_current	Statutory	liabilities_misc_current
	Other Short Term Finances	short_term_loans	Statutory	short_term_loans	Statutory	short_term_loans	Statutory	short_term_loans	Statutory	short_term_loans	Statutory	short_term_loans
	Current Liabilities	liabilities_current	Statutory	liabilities_current	Statutory	liabilities_current	Statutory	liabilities_current	Statutory	liabilities_current	Statutory	liabilities_current
	Contingent Liability	contingent_liability	Statutory	contingent_liability	Statutory	contingent_liability	Statutory	contingent_liability	Statutory	contingent_liability	Statutory	contingent_liability
	Other Long Term Finances	lt_loans	Statutory	lt_loans	Statutory	lt_loans	Statutory	lt_loans	Statutory	lt_loans	Statutory	lt_loans
	Total Long Term Liabilities	liabilities_lt	Statutory	liabilities_lt	Statutory	liabilities_lt	Statutory	liabilities_lt	Statutory	liabilities_lt	Statutory	liabilities_lt
	Total Liabilities	liabilities_total	Statutory	liabilities_total	Statutory	liabilities_total	Statutory	liabilities_total	Statutory	liabilities_total	Statutory	liabilities_total
	Net Assets	assets_net	Statutory	assets_net	Statutory	assets_net	Statutory	assets_net	Statutory	assets_net	Statutory	assets_net
	Equity Paid Up	paid_up_equity	Statutory	paid_up_equity	Statutory	paid_up_equity	Statutory	paid_up_equity	Statutory	paid_up_equity	Statutory	paid_up_equity
	Revaluation Reserve	revaluation_reserve	Statutory	revaluation_reserve	Statutory	revaluation_reserve	Statutory	revaluation_reserve	Statutory	revaluation_reserve	Statutory	revaluation_reserve
	Sundry Reserves	sundry_reserves	Statutory	sundry_reserves	Statutory	sundry_reserves	Statutory	sundry_reserves	Statutory	sundry_reserves	Statutory	sundry_reserves
	Profit and Loss Account Reserve	pandl_account_reserve	Statutory	pandl_account_reserve	Statutory	pandl_account_reserve	Statutory	pandl_account_reserve	Statutory	pandl_account_reserve	Statutory	pandl_account_reserve
	Shareholder Funds	shareholder_funds	Statutory	shareholder_funds	Statutory	shareholder_funds	Statutory	shareholder_funds	Statutory	shareholder_funds	Statutory	shareholder_funds
Other	Wages & Salaries	wages	Statutory	wages	GAAP	employee_costs	IFRS	employee_costs	Statutory	wages	Statutory	wages
	Directors Emoluments	directors_emoluments	Statutory	directors_emoluments	GAAP	directors_remuneration	IFRS	directors_remuneration	Statutory	directors_emoluments	Statutory	directors_emoluments
	Audit Fees	audit_fees	Statutory	audit_fees	GAAP	statutory_audit_costs	IFRS	statutory_audit_costs	Statutory	audit_fees	Statutory	audit_fees
	Depreciation	depreciation	Statutory	depreciation	GAAP	depreciation_of_tangibles	IFRS	depreciation_of_tangibles	Statutory	depreciation	Statutory	depreciation
	Amortisation of Intangibles	amortisation_of_intangibles	null	null	GAAP	amortisation_of_intangibles	IFRS	amortisation_of_intangibles	INS	amortisation_of_intangibles	null	null
	EBITDA	ebitda	null	null	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Working Capital	working_capital	Statutory	working_capital	Statutory	working_capital	Statutory	working_capital	Statutory	working_capital	Statutory	working_capital
	Bank Overdraft & Long Term Loans	bank_overdraft_lt_loans	Statutory	bank_overdraft_lt_loans	Statutory	bank_overdraft_lt_loans	Statutory	bank_overdraft_lt_loans	Statutory	bank_overdraft_lt_loans	Statutory	bank_overdraft_lt_loans
	Capital Employed	capital_employed	Statutory	capital_employed	Statutory	capital_employed	Statutory	capital_employed	Statutory	capital_employed	Statutory	capital_employed
	Net Cash Flow from Operations	operations_net_cashflow	Statutory	operations_net_cashflow	Statutory	operations_net_cashflow	Statutory	operations_net_cashflow	Statutory	operations_net_cashflow	Statutory	operations_net_cashflow
	Net Cash Flow before Financing	net_cashflow_before_financing	Statutory	net_cashflow_before_financing	Statutory	net_cashflow_before_financing	Statutory	net_cashflow_before_financing	Statutory	net_cashflow_before_financing	Statutory	net_cashflow_before_financ
	Net Cash Flow from Financing	net_cashflow_from_financing	Statutory	net_cashflow_from_financing	Statutory	net_cashflow_from_financing	Statutory	net_cashflow_from_financing	Statutory	net_cashflow_from_financing	Statutory	net_cashflow_from_financi
	Export	exports	Statutory	exports	Statutory	exports	Statutory	exports	Statutory	exports	Statutory	exports
	Interest Payments	interest_payments	Statutory	interest_payments	Statutory	interest_payments	Statutory	interest_payments	Statutory	interest_payments	Statutory	interest_payments
	Increase In Cash	increase_in_cash	Statutory	increase_in_cash	Statutory	increase_in_cash	Statutory	increase_in_cash	Statutory	increase_in_cash	Statutory	increase_in_cash
Audit	Name of Auditors	auditors	Statutory	auditors	Statutory	auditors	Statutory	auditors	Statutory	auditors	Statutory	auditors
	Joint Auditors (If Applicable)	joint_auditors	Statutory	joint_auditors	Statutory	joint_auditors	Statutory	joint_auditors	Statutory	joint_auditors	Statutory	joint_auditors
	Name of Solicitors	solicitors	Statutory	solicitors	Statutory	solicitors	Statutory	solicitors	Statutory	solicitors	Statutory	solicitors
	Name of Accountants	accountants	Statutory	accountants	Statutory	accountants	Statutory	accountants	Statutory	accountants	Statutory	accountants
	Audit Qualification	qualification_code	Statutory	qualification_code	Statutory	qualification_code	Statutory	qualification_code	Statutory	qualification_code	Statutory	qualification_code
Ratios	Cash to Current Liabilities Ratio	cash_to_current_liabilities_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Cash to Total Assets (%)	cash_to_total_assets_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Current Ratio	current_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Gearing (%)	gearing	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Gross Margin (%)	gross_margin_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Inventory Turnover Ratio	inventory_turnover_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null
	Liquidity Ratio	liquidity_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Net Margin (%)	profit_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Return on Assets (%)	return_on_assets_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Return on Capital Employed (%)	return_on_capital_employed	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Debtor Days	debtor_days	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Debt to Capital (%)	debt_to_capital_ratio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Operating Profit Margin (%)	operating_profit_margin	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null		null	
	EBITDA Margin (%)	ebitda_margin	null	null	DERIVED	calculation	DERIVED	calculation	null		null	
	Pre-tax Profit Margin (%)	pre_tax_profit_margin	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Return on Equity (%)	return_on_equity	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Days Inventory Outstanding (days)	dio	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Days Sales Outstanding (days)	dso	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Days Payable Outstanding (days)	dpo	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Cash Conversion Cycle (days)	ccc	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	null	null
	Revenue per Employee (annualized)	revenue_per_employee	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation
	Human Capital Value Added (annualized)	human_capital_value_added	DERIVED	calculation	DERIVED	calculation	DERIVED	calculation	null	null	DERIVED	calculation

Identifying the previous set of accounts



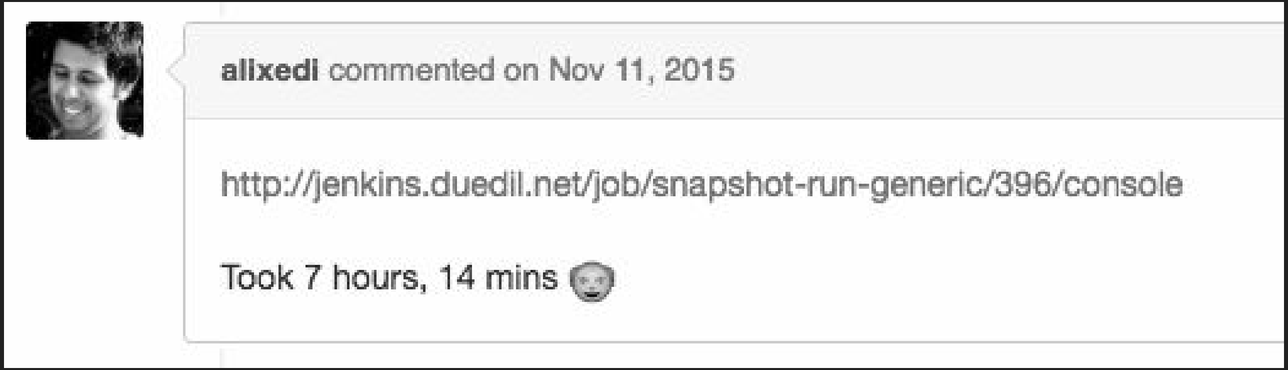
This allows:

- Correct calculation of deltas and percentage deltas even if there are multiple accounts filed for largely overlapping time periods;
- Calculation of average balance sheet values over an accounting period for use in certain ratios and KPIs.

#Bonus
Nothing beats awesome specs

www.duedil.com/careers

Questions



alixedi commented on Nov 11, 2015

http://jenkins.duedil.net/job/snapshot-run-generic/396/console

Took 7 hours, 14 mins

Hello: @alixedi, GitHub: alixedi, Essays: alixedi.github.io