new codebase who dis?

Samuel Taylor
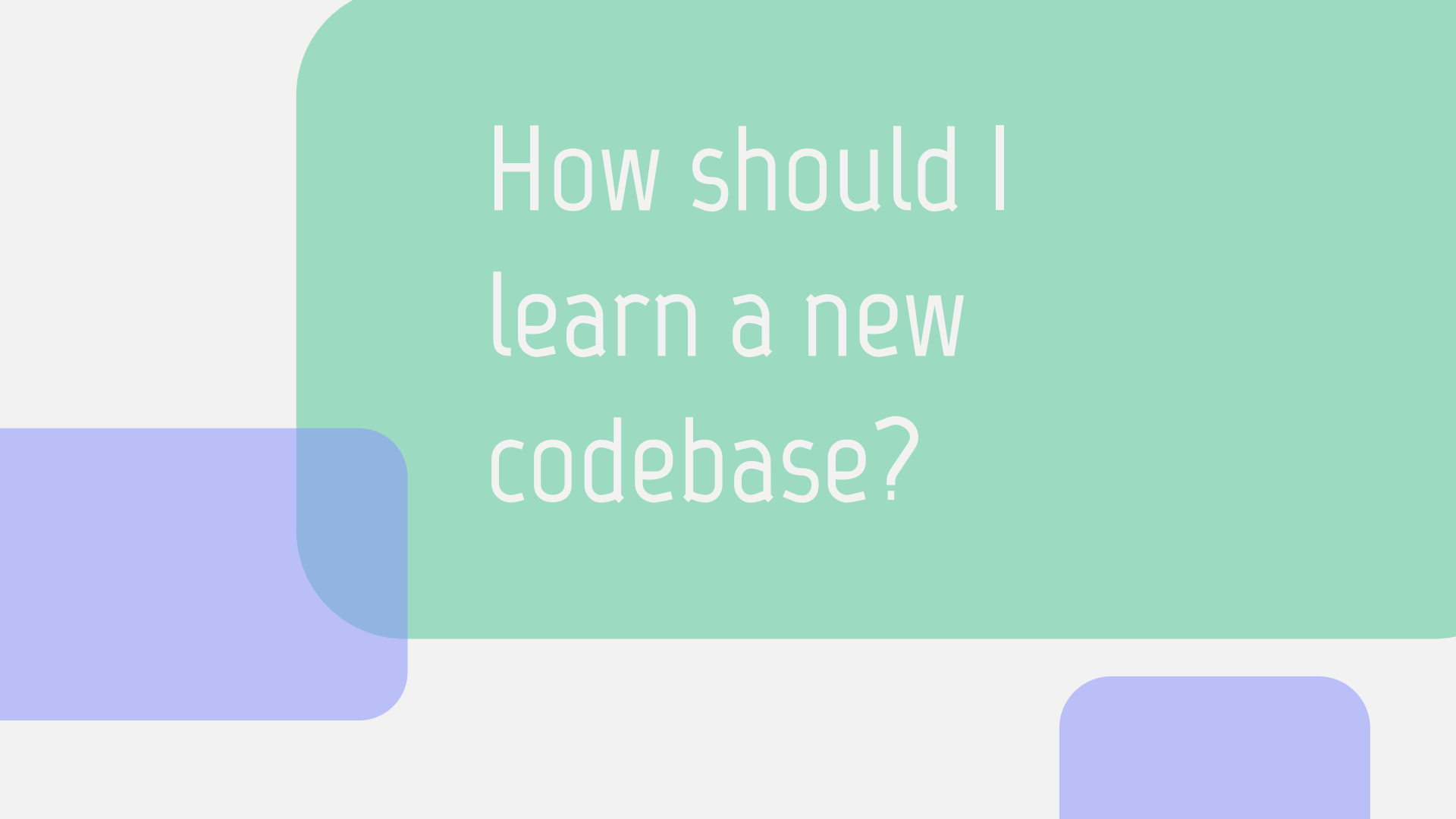Sr. AI Engineer
Acrisure Technology Group

PIZZA
HOT & DELICIOUS

$

☐ PEPPERONI   ☐ CHEESE

drewsoulman

PIZZA
HOT & DELICIOUS

$

☐ PEPPERONI   ☐ CHEESE

drewsoulman

How should I
learn a new
codebase?

PROCESS

DAY 1

00

01

02

03

04

INTRO
(you are here)

MINDSET

TOOLS

Day 1

Three steps to success

Paying attention while setting up the dev environment will reap dividends

- Understand what services you're running
- Understand interservice dependencies
- Take specific notes
- Run the tests, then break them

Senior members of the team should help you understand the architecture

- Look for a document/diagram
- Ask questions
  - What repos do we own?
  - How does a feature get from my laptop to prod?
  - What vendors/API's do we rely on?
- Create or update the architecture document

You are doomed if you don't understand the business

- Understand mission, products, goals
- Figure out your team's contribution
    - How can we impact company goals?
    - Who gets mad if our code breaks horrifically?

Mindset

Learn by doing

"Ultralearning is like a superpower in our competitive economy. Read this book! It will change your life."
—CAL NEWPORT, author of *Digital Minimalism* and *Deep Work*

ULTRALEARNING

Master Hard Skills,
Outsmart the Competition, and
Accelerate Your Career

Scott H. Young

Foreword by JAMES CLEAR, author of *Atomic Habits*

Antonello Calamea

The easiest way to learn directly is to simply spend a lot of time doing the thing you want to become good at.
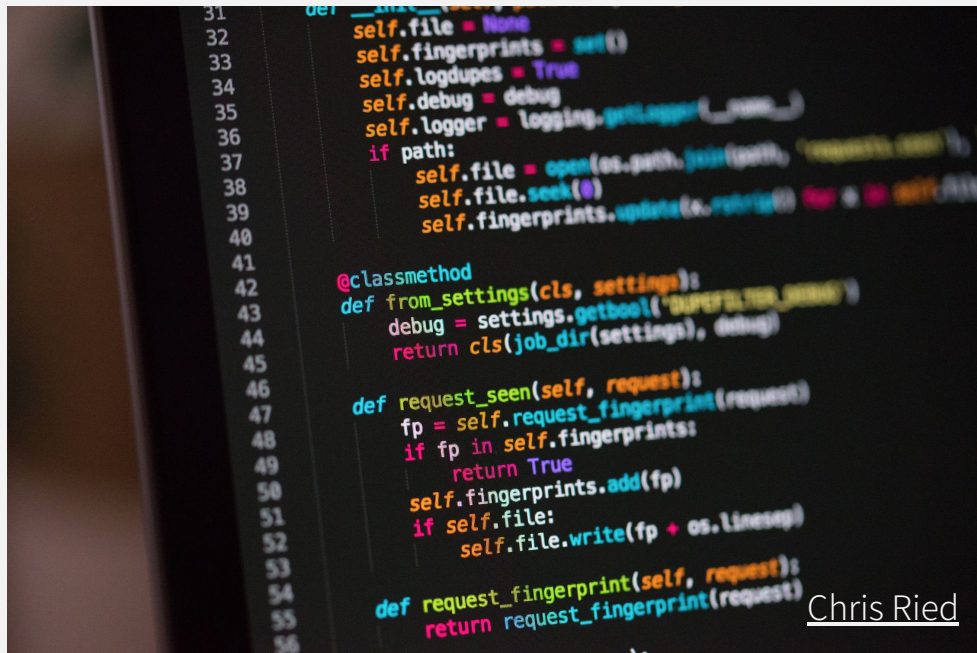
Scott Young,
Ultralearning

Deep understanding comes by making an impact

- DO: Understand code well enough to make a change
- DON'T: Read every line

Reading a book is very different from reading code

Blaz Photo

Chris Ried

Your knowledge will grow recursively

1. Understand what the services do
2. Identify what service you need to modify
3. Understand modules of that service
4. Within a given module, know the classes
5. …
6. Understand an individual line of code

"Chunking" is the brain's ability to bind detailed information into a concept that is easy to remember

Michael Benz

Think in terms of code paths and data flows

GET /puppy

getPuppy()

PuppyManager
.getRandomPuppy()

DBAccessor.
runQuery()
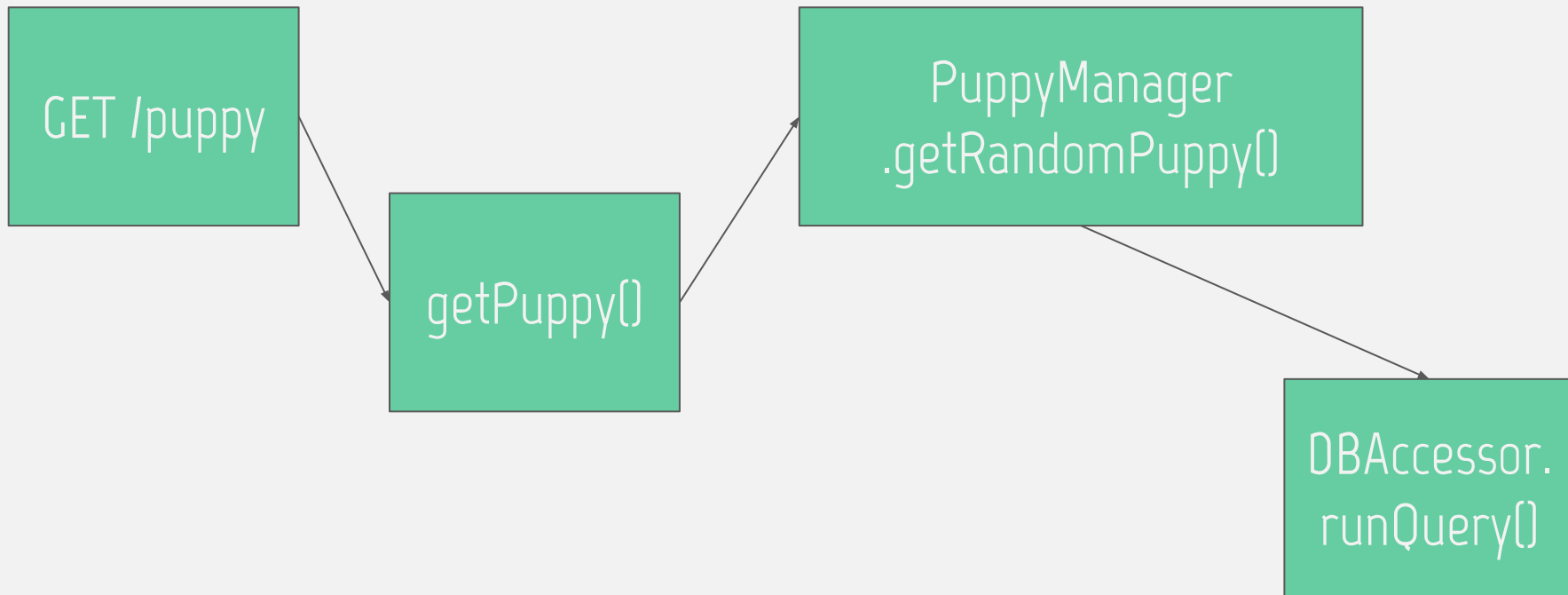
Think in terms of code paths and data flows
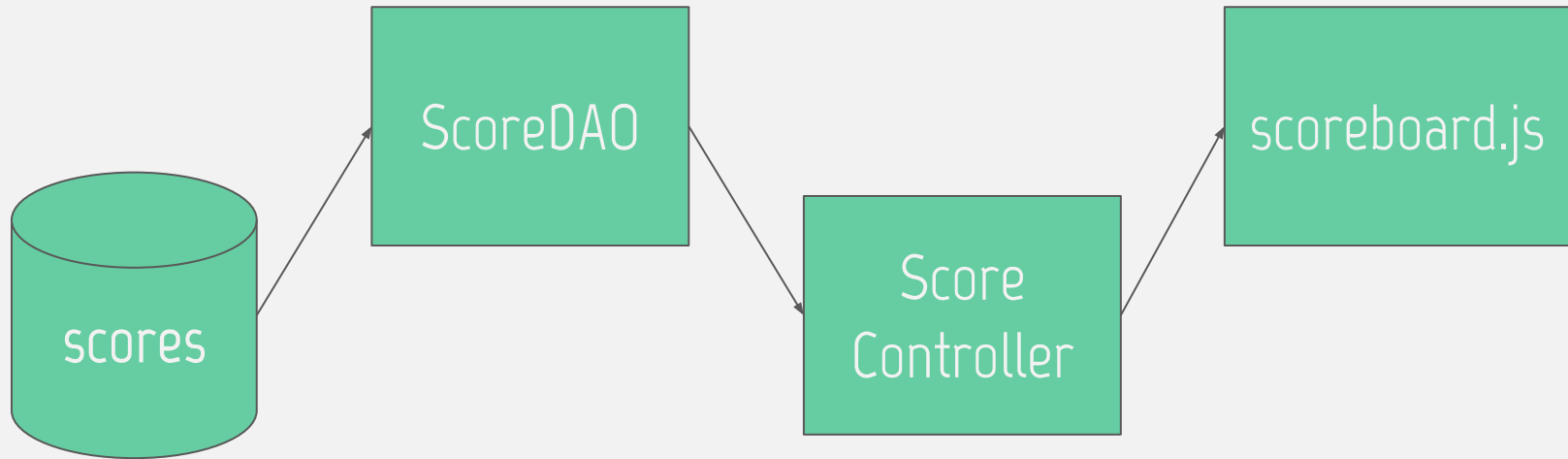
Process

Use the scientific method

01

Find relevant code

01 Find relevant code

02 Form a hypothesis
about what you need to
change

01 Find relevant code

02 Form a hypothesis
about what you need to change

Test your hypothesis
after making the change. Return to 02 if you were wrong. **03**

01 Find relevant code

02 Form a hypothesis
about what you need to change

Test your hypothesis
after making the change. Return to 02 if you were wrong. **03**

Improve quality **04**

Tools

Finding relevant code

Run the code

- Understand what exists
- Use the debugger

Finding relevant code

Search the project

- JIRA, Asana, Pivotal Tracker
- GitHub, GitLab issues

Dashboards ⌄ Projects ⌄ Issues ⌄

Search 🔍 ❓ Log In

New search «

Find filters

FILTERS

My open issues

Reported by me

All issues

Open issues

Done issues

Viewed recently

Created recently

Resolved recently

Updated recently

FAVOURITE FILTERS

You must be logged in to view favourite filters.

Search   Save as

📤 Export ⌄   ⚙ Tools ⌄

Project... ⌄   Type: All ⌄   Status:... ⌄   Assign... ⌄   [datetime]   More ⌄   **Search**   Advanced   ⊟⚙ ⌄

Order by ⌄   ↑

ARROW-7528
[Python] The pandas.datetime class (...

BEAM-4723
Enhance Datetime*Expression Dateti...

BEAM-5510
Records including datetime to be sav...

PIG-3341
Strict datetime parsing and improve ...

WICKET-466
[datetime] Make DateConverter focu...

LUCENENET-231
System.DateTime.Now.Millisecond bug

DAFFODIL-105
DateTime - Text

SPARK-22966
Spark SQL should handle Python UD...

AXIS-1481
No deserializer for dateTime

ARROW-9755
[Python] pyarrow deserialize return d...

ARROW-8066
[Python] Specify behavior for conver...

ASTERIXDB-1430

Pig / PIG-3341

4 of 3893 ⌃ ⌄

Strict datetime parsing and improve performance of loading datetime values

📤 Export ⌄

⌄ **Details**

Type: ⬆ Improvement

Status: **CLOSED**

Priority: ⚠ Major

Resolution: Fixed

Affects Version/s: 0.11.1

Fix Version/s: 0.12.0, 0.11.2

Component/s: impl

Labels: None

⌄ **Description**

The performance of loading datetime values can be improved by about 25% by moving a single line in ToDate.java:

public static DateTimeZone extractDateTimeZone(String dtStr) {
Pattern pattern = Pattern.compile("
(Z|(?<=(T[0-9\\.:]
{0,12}))((\\+|-)
d{2}(:?

⌄ **People**

Assignee:
⬤ Rohini Palaniswamy

Reporter:
⬤ pat chan

Votes:
0 Vote for this issue

Watchers:
4 Start watching this issue

⌄ **Dates**

Created:
31/May/13 06:11

Updated:
01/Oct/19 22:10

Resolved:
13/Jun/13 13:07

Enable local test runner for Kubeflow Pipelines #1138

New issue

⊙ Open   neuromage opened this issue on Apr 11, 2019 · 15 comments

---

neuromage commented on Apr 11, 2019   Collaborator   ···

Today, there is no way to run a KFP pipeline on a local machine (outside of running a mini Kubernetes cluster locally). This makes testing pipelines for user difficult as they always have to go through the steps of uploading to a cluster to run even small tests on the correctness of the pipeline.

To solve this problem, it would be nice if KFP offers a way to run pipelines locally. Such a mode does not have to have full fidelity with what's possible in a cluster run with Argo to begin with. A simple approach of using `docker run` to run each step sequentially, with a mounted local volume for passing parameters, will go a long way towards satisfying most local-run use-cases.

As a suggestion, we can modify the pipeline decorator to enable this behaviour when the user adds a keyword like `test = True` , i.e.

```
@dsl.Pipeline(
    name = "..",
    test = True,
    ...
```

👍 7

🏷 🔲 **neuromage** added the `feature_request` label on Apr 11, 2019

🔳 **issue-label-bot** `bot` commented on Apr 11, 2019   ···

Issue-Label Bot is automatically applying the label `feature_request` to this issue, with a confidence of 0.98. Please mark this comment with 👍 or 👎 to give our bot feedback!

Links: dashboard, app homepage and code for this bot.

Assignees

👤 Ark-kun

👤 rmgogogo

Labels

area/sdk/dsl   kind/feature

lifecycle/frozen

needs investigation   priority/p2

Projects

▤ KFP Post 1.0 Backlog

▬▬▬▬▬▬▬▬▬▬▬

Backlog

Milestone

No milestone

Linked pull requests

Successfully merging a pull request may close this issue.

None yet

9 participants

Finding relevant code

Search the code

- Silver Searcher (ag), ripgrep

```
$ ag phoneme
src/app.py
22:def word_to_phonemes(word):
75:logger.info('Loading phoneme data')
107:    target = word_to_phonemes(input_string)
115:        word_phonemes = [word_to_phonemes(w) for w in words]
117:        scored_idioms.extend((lev_dist(target, word_phoneme) / len(word_phoneme)
 replace_word(idiom, ndx, input_string))
118:                            for ndx, word_phoneme in enumerate(word_phonemes)
119:                            if word_phoneme is not None)

src/data/cmudict
91317:PHONEME  F OW1 N IY0 M
91318:PHONEMES  F OW1 N IY0 M Z
```

Finding relevant code

Search the code

- Silver Searcher (ag), ripgrep
- OpenGrok, SourceGraph
- GitHub, GitLab org search

{0 Search ✕ | λ Ord a => [a] -> [a] - ✕ | OH Open Hub Projects ✕ | OH Results : cdef:SortUti ✕

Gerbrand

⟨ ⟩ ↻ ⌂ 🗋 localhost:8080/source/search?q=&defs=*Util*&refs=&p ☆ 🔒 ▽ ● 🗔 ● ◈ 🔥 ≋ 🗹 ● 🎯 ● ● JB 🐝 ≡

▦ Apps ★ Bookmarks 🗨 Groups ▭ RIGD-Loxia ▭ RWE 🔴 reddit 🗋 Top Open-Source ● WhatsApp Web 🐍 Python » ▭ Andere bladwijzers

{OpenGrok

Home   Sort by: last modified time | **relevance** | path

Full Search [   ]

Definition [*Util*   ]

Symbol [   ]

File Path [   ]

History [   ]

Type [Java ▾]

In Project(s)   [select all] [invert selection]

```
Transip
VanDieijen
angularplay
blikvanger
exportiphoto
gerbrand-ict.nl
twistedmind.nu
```

[**Search**] [**Clear**] [Help]

Searched **+defs:*Util* +type:java** (Results **1 - 5** of **5**) sorted by relevance

/Engelvaart/engelvaart/VanDieijenCommon/src/com/vandieijen/common/

HAD	FileUtil.java	16	`public final class `**`FileUtil`**` { ` *class*
HAD	FormUtil.java	19	`final public class `**`FormUtil`**` { ` *class*
HAD	GUIResourceUtilizer.java	48	`final public class `**`GUIResourceUtilizer`**` { ` *class*
		82	`public `**`GUIResourceUtilizer`**`(String name) { ` *method in class:GUIResourceUtilizer*
		86	`public `**`GUIResourceUtilizer`**`(ResourceBundle resources) { ` *method in class:GUIResourceUtilizer*
HAD	Util.java	22	`public class `**`Util`**` { ` *class*
		27	`public `**`Util`**`() { ` *method in class:Util*

Gerbrand

Understanding code

- IDE
    - (I like JetBrains' Cmd+Click)
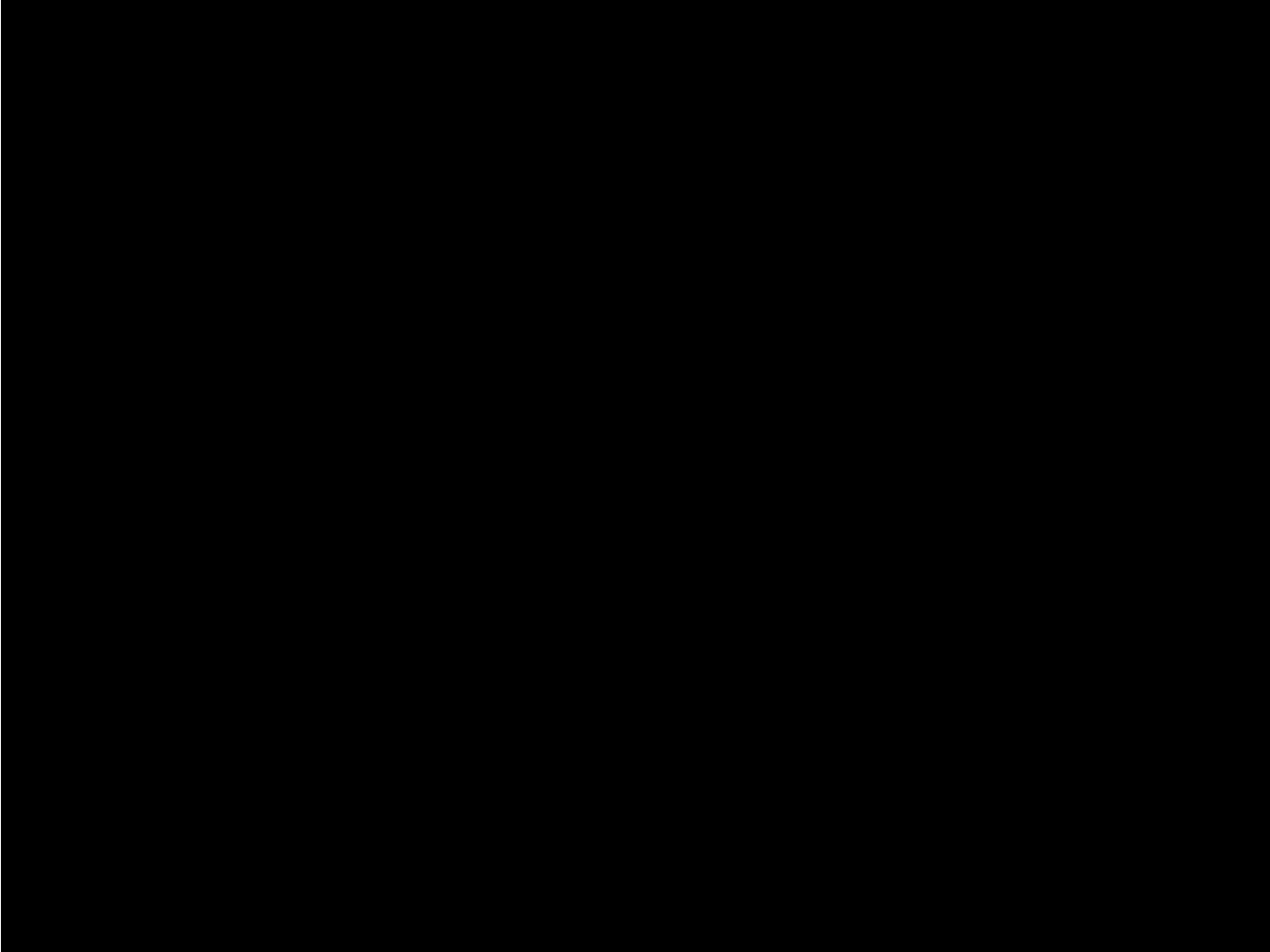
Understanding code

- IDE
  - (I like JetBrains' Cmd+Click)
- Remember to create "chunks"

Understanding code

- IDE
  - (I like JetBrains' Cmd+Click)
- Remember to create "chunks"
- Notes, diagrams

Understanding code

- IDE
    - (I like JetBrains' Cmd+Click)
- Remember to create "chunks"
- Notes, diagrams
- Ask for help (use git blame)

```python
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 55)        result = [value_counts_without_nan, distinct_count_with_nan]
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 56)
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 57)        if data.name is not None:
401ca63d (romainx                2018-01-04 22:20:40 +0100 58)            _VALUE_COUNTS_MEMO[data.name] = result
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 59)
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 60)        return result
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 61)
401ca63d (romainx                2018-01-04 22:20:40 +0100 62) _MEMO = {}
9e5848d2 (romainx                2018-01-01 11:05:19 +0100 63) def get_vartype(data):
401ca63d (romainx                2018-01-04 22:20:40 +0100 64)    """Infer the type of a variable (technically a Series).
bdb43228 (romainx                2018-01-01 18:17:02 +0100 65)
bdb43228 (romainx                2018-01-01 18:17:02 +0100 66)    The types supported are split in standard types and special types.
6a452f7f (jos                    2016-01-24 14:41:16 -0500 67)
9e5848d2 (romainx                2018-01-01 11:05:19 +0100 68)    Standard types:
bdb43228 (romainx                2018-01-01 18:17:02 +0100 69)        * Categorical (`TYPE_CAT`): the default type if no other one can be determined
bdb43228 (romainx                2018-01-01 18:17:02 +0100 70)        * Numerical (`TYPE_NUM`): if it contains numbers
bdb43228 (romainx                2018-01-01 18:17:02 +0100 71)        * Boolean (`TYPE_BOOL`): at this time only detected if it contains boolean value
s, see todo
bdb43228 (romainx                2018-01-01 18:17:02 +0100 72)        * Date (`TYPE_DATE`): if it contains datetime
6a452f7f (jos                    2016-01-24 14:41:16 -0500 73)
9e5848d2 (romainx                2018-01-01 11:05:19 +0100 74)    Special types:
bdb43228 (romainx                2018-01-01 18:17:02 +0100 75)        * Constant (`S_TYPE_CONST`): if all values in the variable are equal
bdb43228 (romainx                2018-01-01 18:17:02 +0100 76)        * Unique (`S_TYPE_UNIQUE`): if all values in the variable are different
79e71b98 (Conrado Quilles Gomes 2018-01-02 18:33:55 -0200 77)        * Unsupported (`S_TYPE_UNSUPPORTED`): if the variable is unsupported
e4a0998e (jos                    2016-08-14 18:49:36 -0400 78)
401ca63d (romainx                2018-01-04 22:20:40 +0100 79)     The result is cached by column name in a global variable to avoid recomputing.
401ca63d (romainx                2018-01-04 22:20:40 +0100 80)
bdb43228 (romainx                2018-01-01 18:17:02 +0100 81)    Parameters
bdb43228 (romainx                2018-01-01 18:17:02 +0100 82)    ----------
bdb43228 (romainx                2018-01-01 18:17:02 +0100 83)    data : Series
bdb43228 (romainx                2018-01-01 18:17:02 +0100 84)        The data type of the Series.
85e622a6 (romainx                2016-05-22 01:23:48 +0200 85)
bdb43228 (romainx                2018-01-01 18:17:02 +0100 86)    Returns
bdb43228 (romainx                2018-01-01 18:17:02 +0100 87)    -------
bdb43228 (romainx                2018-01-01 18:17:02 +0100 88)    str
bdb43228 (romainx                2018-01-01 18:17:02 +0100 89)        The data type of the Series.
85e622a6 (romainx                2016-05-22 01:23:48 +0200 90)
bdb43228 (romainx                2018-01-01 18:17:02 +0100 91)    Notes
bdb43228 (romainx                2018-01-01 18:17:02 +0100 92)    ----
0762fd54 (Conrado Quilles Gomes 2018-01-02 00:44:47 -0200 93)        * Should improve verification when a categorical or numeric field has 3 values,
it could be a categorical field
0762fd54 (Conrado Quilles Gomes 2018-01-02 00:44:47 -0200 94)        or just a boolean with NaN values
9e5848d2 (romainx                2018-01-01 11:05:19 +0100 95)        * #72: Numeric with low Distinct count should be treated as "Categorical"
e4a0998e (jos                    2016-08-14 18:49:36 -0400 96)    """
:
```

Working with libraries

- Library docs
- StackOverflow
- GitHub search

Repositories	0
Code	1K
Commits	15
Issues	143
Discussions (Beta)	0
Packages	0
Marketplace	0
Topics	0
Wikis	0
Users	0

Languages

Java	643
PHP	369
Scala	61
Clojure	13
HTML	13
Markdown	12
MATLAB	9
Kotlin	6
Java Server Pages	4
Groovy	3

Advanced search   Cheat sheet

1,200 code results

Sort: Best match ▾

HardCoreWare/mtvpro

vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

```
30     * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
         data.github_repos.commits` LIMIT 100');
31     * ```
32     */
33    class QueryJobConfiguration implements JobConfigurationInterface
34    {
35        use JobConfigurationTrait;
       ...
67     * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
68     * be set.
69     *
70     * Example:
71     * ```
72     * $query->allowLargeResults(true);
```

● PHP   Showing the top two matches   Last indexed on Oct 24, 2018

Danden7/rekrutacja2

vendor/google/cloud/src/BigQuery/QueryJobConfiguration.php

```
30     * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
         data.github_repos.commits` LIMIT 100');
31     * ```
32     */
33    class QueryJobConfiguration implements JobConfigurationInterface
34    {
35        use JobConfigurationTrait;
       ...
60     * performed with legacy SQL dialect and requires a
61     * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
62     * be set.
```

● PHP   Showing the top two matches   Last indexed on Nov 8, 2019

Clicker1988/administration

admin/application/vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

```
30     * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
```

QueryJobConfiguration /

Pull requests   Issues   Marketplace   Explore

1,200 code results

Sort: Best match ▾

Repositories 0
Code 1K
Commits 15
Issues 143
Discussions (Beta) 0
Packages 0
Marketplace 0
Topics 0
Wikis 0
Users 0

Languages
Java 643
PHP 369
Scala 61
Clojure 13
HTML 13
Markdown 12
MATLAB 9
Kotlin 6
Java Server Pages 4
Groovy 3

Advanced search   Cheat sheet

HardCoreWare/mtvpro
vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

```
30      * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
           data.github_repos.commits` LIMIT 100');
31      * ```
32      */
33    class QueryJobConfiguration implements JobConfigurationInterface
34    {
35        use JobConfigurationTrait;
      ...
67      * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
68      * be set.
69      *
70      * Example:
71      * ```
72      * $query->allowLargeResults(true);
```
● PHP   Showing the top two matches   Last indexed on Oct 24, 2018

Danden7/rekrutacja2
vendor/google/cloud/src/BigQuery/QueryJobConfiguration.php

```
30      * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
           data.github_repos.commits` LIMIT 100');
31      * ```
32      */
33    class QueryJobConfiguration implements JobConfigurationInterface
34    {
35        use JobConfigurationTrait;
      ...
60      * performed with legacy SQL dialect and requires a
61      * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
62      * be set.
```
● PHP   Showing the top two matches   Last indexed on Nov 8, 2019

Clicker1988/administration
admin/application/vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

Repositories	0
Code	1K
Commits	15
Issues	143
Discussions (Beta)	0
Packages	0
Marketplace	0
Topics	0
Wikis	0
Users	0

Languages

Java	643
PHP	369
Scala	61
Clojure	13
HTML	13
Markdown	12
MATLAB	9
Kotlin	6
Java Server Pages	4
Groovy	3

Advanced search   Cheat sheet

1,200 code results

Sort: Best match ▾

HardCoreWare/mtvpro
vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

```
30       * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
            data.github_repos.commits` LIMIT 100');
31       * ```
32       */
33      class QueryJobConfiguration implements JobConfigurationInterface
34      {
35          use JobConfigurationTrait;
        ...
67       * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
68       * be set.
69       *
70       * Example:
71       * ```
72       * $query->allowLargeResults(true);
```

● PHP   Showing the top two matches   Last indexed on Oct 24, 2018

Danden7/rekrutacja2
vendor/google/cloud/src/BigQuery/QueryJobConfiguration.php

```
30       * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
            data.github_repos.commits` LIMIT 100');
31       * ```
32       */
33      class QueryJobConfiguration implements JobConfigurationInterface
34      {
35          use JobConfigurationTrait;
        ...
60       * performed with legacy SQL dialect and requires a
61       * {@see Google\Cloud\BigQuery\QueryJobConfiguration::destinationTable()} to
62       * be set.
```

● PHP   Showing the top two matches   Last indexed on Nov 8, 2019

Clicker1988/administration
admin/application/vendor/google/cloud/BigQuery/src/QueryJobConfiguration.php

```
30       * $query = $bigQuery->query('SELECT commit FROM `bigquery-public-
```

643 code results

Sort: Recently indexed ▾

Repositories	0
Code	643
Commits	15
Issues	133
Discussions (Beta)	0
Packages	0
Marketplace	0
Topics	0
Wikis	0
Users	0

all-of-us/workbench

api/src/main/java/org/pmiops/workbench/billing/FreeTierBillingService.java

```
   4    import com.google.cloud.bigquery.QueryJobConfiguration;
   5    import com.google.common.collect.Sets;
   6    import com.google.common.math.DoubleMath;
  ...
 248            .collect(Collectors.toMap(DbWorkspace::getWorkspaceNamespace,
        Function.identity()));
 249
 250        final QueryJobConfiguration queryConfig =
 251            QueryJobConfiguration.newBuilder(
```
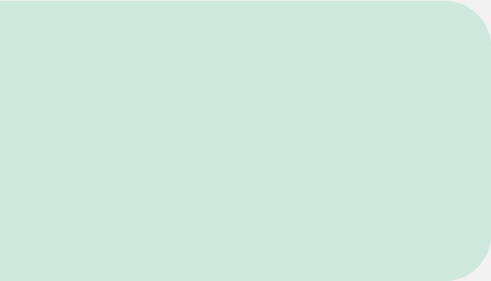
● Java   Showing the top three matches   Last indexed 3 hours ago

GoogleCloudPlatform/professional-services

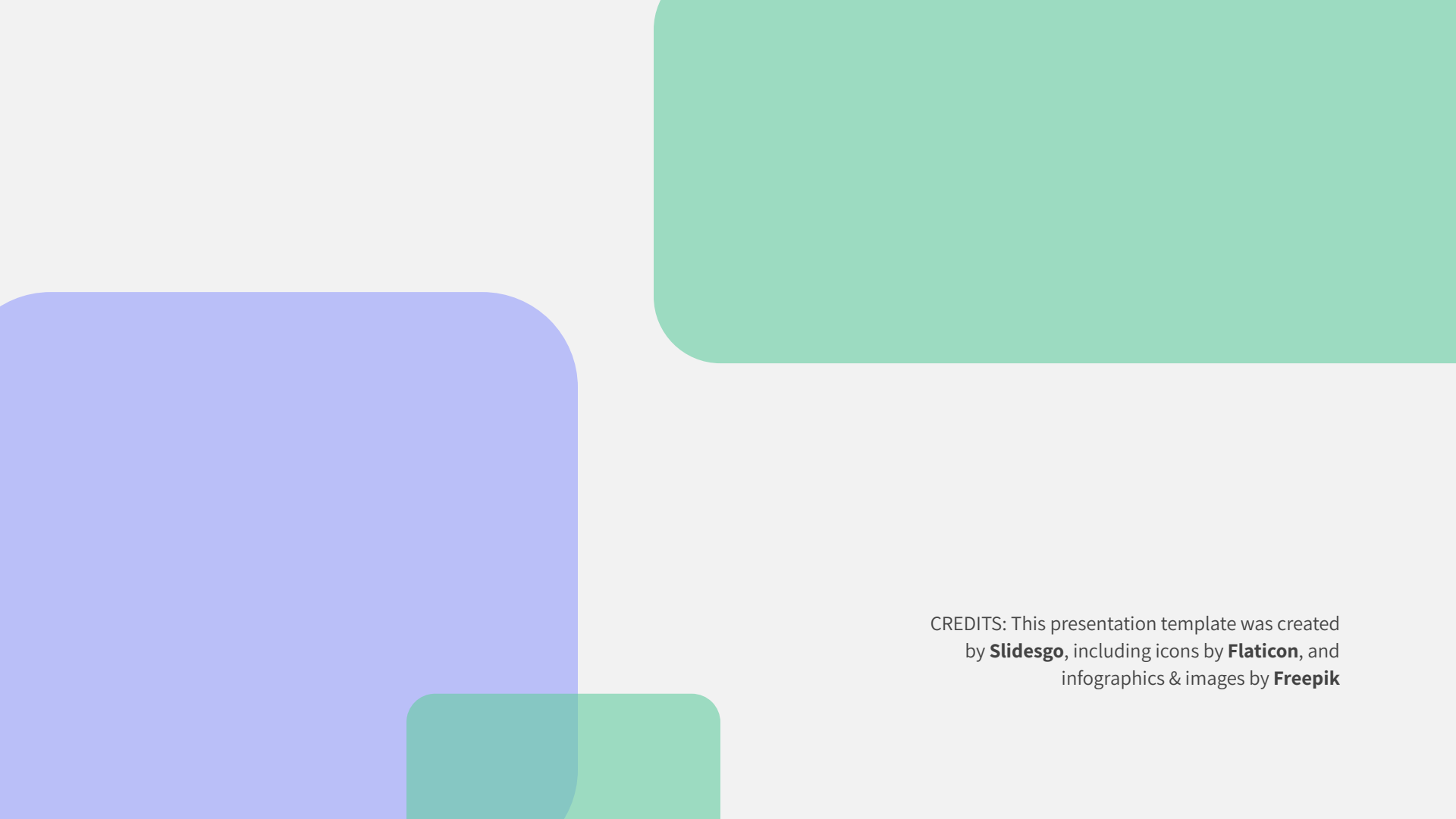**examples/e2e-test-dataflow/src/main/java/util/BQUtils.java**

```
  22    import com.google.cloud.bigquery.QueryJobConfiguration;
  23    import java.util.logging.Logger;
  24
  25    public class BQUtils {
  26      private static final BigQuery bigquery =
```

Languages

Java	✕
PHP	369

- Focus on delivering software
- Provide value to your team
- Decrease cognitive load

Samuel Taylor
@SamuelDataT
sdf@samueltaylor.org