



Technical Skills Shares Group (TSSG) meetup presentation

(Jim Turner TSSG Web Team Manager asked me to share what I have been doing with Cannlytics with regard to web scraping and parsing of url(s)/pdf(s))



<https://www.meetup.com/cannabis-data-science/>



Cannlytics (Cannabis Analytics Open Source Software)

- web scraping with Selenium and BeautifulSoup
- parsing pdf files with PDFplumber

Candace O`Sullivan-Sutherland 8/27/2022

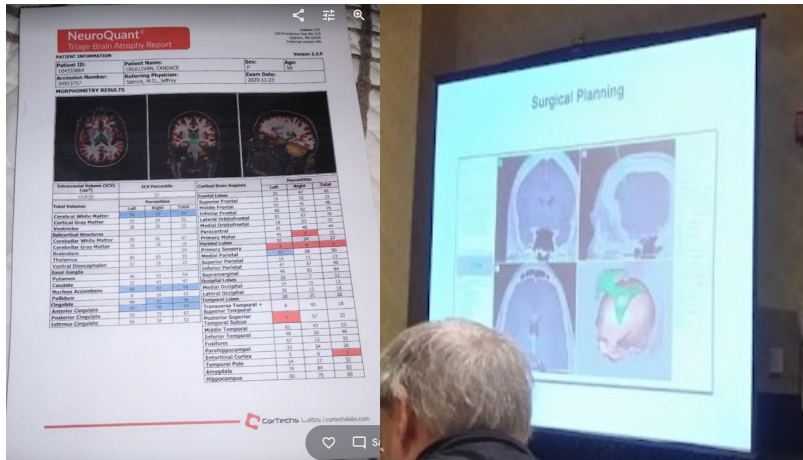
<https://github.com/candy-o>
[2021-present] Software Engineer/Data Scientist @ cannlytics.com
[2019-present] Software Engineer/Data Scientist @ TSSG
[1998-2018] Senior Software Performance Engineer @ Dell Technologies DellEMC (DG CLARiON)
[1995-1998] Senior Software Engineer @ Digital Equipment Corporation, StorageWorks Division



<https://www.meetup.com/technical-skills-share-group/>
<https://www.meetup.com/cannabis-data-science/>
<https://cannlytics.com>
<https://github.com/cannlytics>

Web Scraping and Parsing

- Web scraping and parsing pdf files parsing should be of interest to the TSSG Web group
- For decades “Big Data/Datacenter Storage Performance Root Cause Analysis/Data Science” was my passion previous to switching my Domain Knowledge to “Federally-illegal State-legal Cannabis/Data Science”.
- I am passionate because I believe Cannabis is better for human beings than Pharmaceutical drugs, surgeries,... Plant-based Medical Cannabis replaced **all** my pharmaceutical drugs, procedures and DBS surgery. My Human body receptors achieved homeostasis (healthy) state and my quality of life improved with Cannabis. Cannabis is Medicine, offering both neurological relief and brain damage repair (I am actually coding again!).



Commonly known as: AUGMENTIN	
atorvastatin 10 MG tablet Commonly known as: LIPITOR	atorvastatin 10 mg tablet
baclofen 10 MG tablet Commonly known as: LIORESAL	Take 1 tablet (10 mg total) by mouth 3 (three) times a day.
gabapentin 300 MG capsule Commonly known as: NEURONTIN	gabapentin 300 mg capsule
lorazepam 2 MG tablet Commonly known as: ATIVAN	Take 2 mg by mouth 3 (three) times a day.
oxyCODONE-acetaminophen 5-325 mg per tablet Commonly known as: PERCOCET	oxycodone-acetaminophen 5 mg-325 mg tablet
prazosin 1 MG capsule Commonly known as: MINIPRESS	prazosin 1 mg capsule
prazosin 2 MG capsule Commonly known as: MINIPRESS	prazosin 2 mg capsule
trihexyphenidyl 2 MG tablet Commonly known as: ARTANE	Take 2 mg by mouth 3 (three) times a day with meals.
zolpidem 10 mg tablet Commonly known as: AMBIEN	Take 10 mg by mouth nightly as needed for sleep.

NH_i National Center for Biotechnology Information
Log In

PubMed.gov Advanced Search User Guide

Save Email Send to Display options

Review > J Neural Transm (Vienna). 2019 Jul;127(7):913-924.
doi: 10.1007/s00702-019-02018-h. Epub 2019 May 27.

Evidence for the use of cannabinoids in Parkinson's disease

Carsten Buhmann¹, Tina Mäkinen², Georg Ebersbach³, Florin Gândă⁴*

Affiliations + expand
PMID: 31131434 DOI: 10.1007/s00702-019-02018-h

Abstract

Cannabis and synthetic cannabinoid formulations have now been legally approved in several countries for treatment of patients with Parkinson's disease (PD). Hence, PD patients consult physicians more frequently for prescription of cannabinoids to alleviate symptoms that might not respond well to dopaminergic treatment. Despite increasing volume of cannabis consumption in the field of cannabinoid research and their effect on Parkinson's disease, there is still paucity of sufficient clinical data about the efficacy and safety in PD patients. There is increasing understanding of the endocannabinoid system, and the distribution of cannabinoid receptors in basal ganglia structures might suggest potential benefit on parkinsonian symptoms. Concerning clinical research, only one of to date four conducted randomized placebo-controlled trials showed an effect on motor symptoms with alleviation of levodopa-induced dyskinesia. There are a growing number of uncontrolled trials and case reports that suggest beneficial effects of cannabinoids in PD patients. However, the variety of substances investigated, the varying routes of intake, differing doses and time courses make it difficult to compare data. We here provide an overview of the current literature in this field and discuss a pragmatic approach for the clinical use of cannabinoids in PD.

Keywords: Cannabinoids; Parkinson's disease; THC; Tetrahydrocannabinol.

FULL TEXT LINKS
[SpringerLink](#)

ACTIVE
Cite Favorites

SHARE

PAGE NAVIGATION
Title & authors
Abstract
Similar articles
Cited by
References
Publication types

“Cannlytics” is an Open-source, Decentralized Cannabis Data Science/Analytics organization offering a suite of free software for the Cannabis industry.



Advanced Statistics and Cannabis Data

Automated statistics and the finest curated data for your cannabis OP to push the bounds of what is possible. Open-source, decentralized, and made with ❤️!

[Get Started](#) [Sign Up](#)

Why Cannlytics?



Smart Integrations

We believe that everyone benefits when people are able to study and tinker with their software. With the freedom provided by Cannlytics, users control their software and what it does for them.

[Begin customizing](#)



Analysis Tailored

Cannlytics provides a user-friendly interface to quickly receive samples, perform analyses, collect and review results, and publish certificates of analysis (CoAs). There are also built in logistics, CRM (client relationship management), inventory management, and invoicing tools.

[View capabilities](#)



Community Driven

Built by scientist for scientists. Cannlytics empowers you with control over the development process, resources, and decision making authority. We believe that the Cannlytics community is the best judge of how Cannlytics can be improved, so, we have entrusted the Cannlytics source code with you.

[Join today](#)

Want to explore Cannabis Data Science?



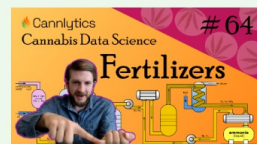
Fertilizer Prices and the US Hemp Canopy | Cannabis Data Science #65

Every piece of the puzzle must be filled in, so we diligently collect all public fertilizer price and hemp yield, harvest, and acreage data under the sun. Please en...

[View more](#)

[Join the Meetup](#)

[Get the code & data](#)



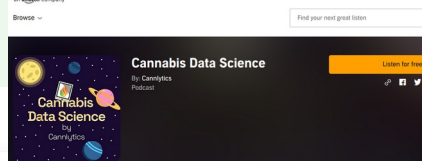
Fertilizers: Costs, Benefits, and Plant Hardiness | Cannabis Data Science #64

What are Nitrogen (N), Phosphorus (P), and Potassium (K) and why should a cannabis cultivator care? This is the exact topic of the day. We follow the money and conn...



Cannabis Consumption: Estimating Consumer Demand | Cannabis Data Science #63

Someone call a plumber! The data dam just burst! The Cannabis Data Science team is all hands on deck serving you in the holy trail of



How can I use Cannlytics?



The Cannlytics Console

If you need a hosted solution, then you can take advantage of the ready to use Cannlytics console for hosting, analysis, and data and file storage. You can manage your lab results and provide your clients with a simple, yet com

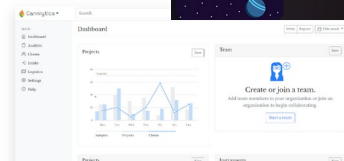
[Create an account](#)



The Cannlytics API

Do you have a technical staff? Then you can utilize the Cannlytics API to programmatically manage your operations and interface with your data. You have all of your data and the power of all laboratory software at your finger tips.

[Read the Docs](#)



```
{
  analyses: [...],
  lab_id: "cann-do",
  sample_name: "Super Easy",
  results: [...],
}
```




Details

🍌 Please join a fun group of data scientists, cannabis enthusiasts, and many more who are interested in applying data science in the cannabis space. Wednesday at 1:20pm PST / 2:20pm MT / 3:20pm CDT / 4:20pm EST at <https://cannlytics.com/meetup>


🍌 At the Cannabis Data Science meetup, you can join like-minded individuals, learn modern data science techniques, and go through hands-on exercises with real cannabis data. This is a great opportunity for anyone in a cannabis-testing lab, anyone working at a cannabis cultivation or processing plant, any cannabis retailer, and even any cannabis researcher or consumer seeking knowledge.

Cannlytics (creator “Keegan” and Cannlytics Team) made with




Cannlytics is built with the love of...

These heros who create, make, and deliver your software solutions and support.




Keegan Skeate
KEO

Keegan is an economist, software developer, and data scientist who cut his chops as a laboratory analyst. Keegan will work tirelessly to deliver you value for value.




Alice Allafort, Ph.D.
Data Scientist

Alice graduated with a Physics Ph.D. from Stanford University and is always happy to work on interesting data science projects.




Jeffrey Miller
Customer Outreach Developer

Jeff is a customer outreach specialist who revolves around creating value and lives and breathes customer satisfaction.



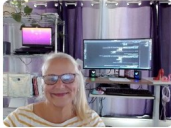
Raymond Bentson
International Communications Specialist / Linguist

Raymond is experienced in agriculture, international trade, and language services. Raymond is also the founder of Advanced Language Liaison.




Michael Pilosov, Ph.D.
Project Manager

Michael is an applied mathematician on a mission to democratize access to quantitative decision-making tools. Michael is here to help you know your data inside and out and interface with your analytics to get the insights that you need.




Candace O'Sullivan-Sutherland
Data Scientist

Candace is a talented data and computer scientist with all the tools in her toolbox to handle the biggest of big data. Candace brings an extensive suite of hardware capabilities to ensure that your data pipelines are pristine.




CannBot
Chief Scientific Robot

CannBot does the dirty work, working meticulously around the clock to automate the boring stuff and do the heavy lifting.



Cat the Caterpillar
Chief Organic Organism

Cat, the data-pipeline caterpillar, ensures that your data is clean, organized, and gets to where you need it.

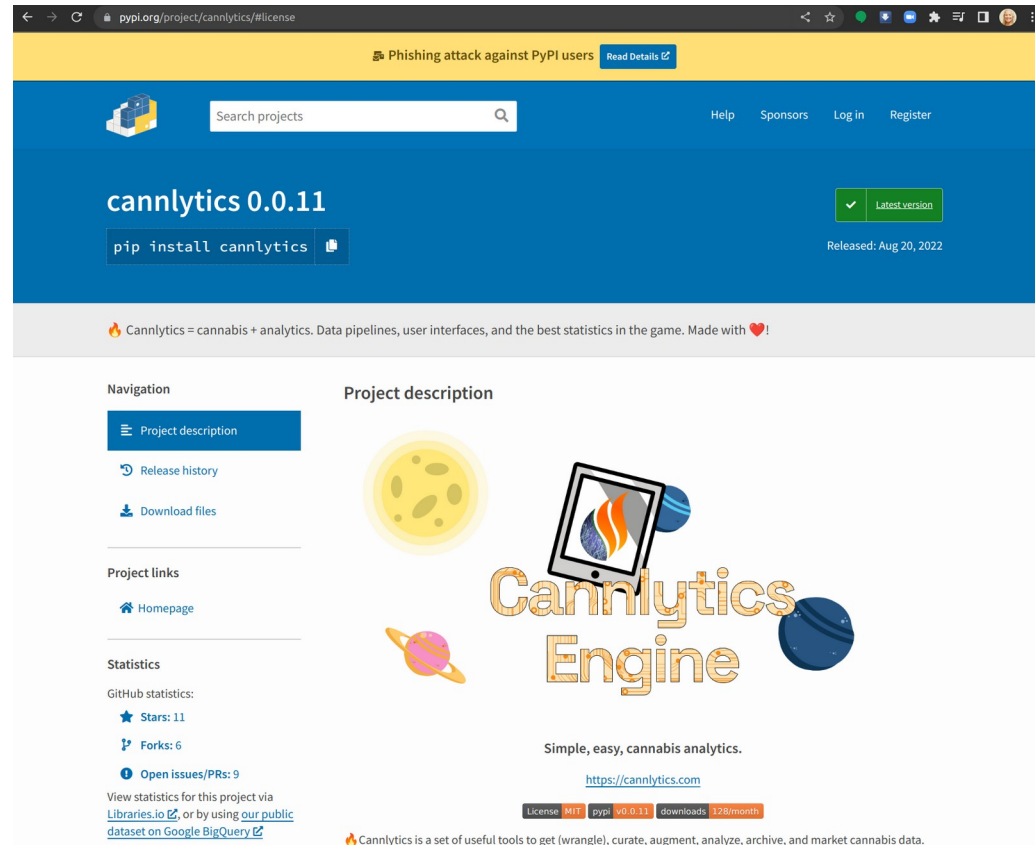


Interested in joining?

Do you have the chops and / or passion to help deliver industry-leading analytics, software, and support to everyone in the cannabis industry? Then please email the team at dev@cannlytics.com and explain who, what, where, why, when, and how you can contribute to Cannlytics.

[Join today →](#)

pipx.org `cannlytics engine` is the heart of cannlytics



The screenshot shows the PyPI project page for 'cannlytics'. At the top, there's a yellow banner about a phishing attack. Below that, the project name 'cannlytics 0.0.11' is displayed in large white text on a blue background. A green button indicates it's the 'Latest version', and the release date 'Released: Aug 20, 2022' is shown. A search bar and navigation links (Help, Sponsors, Log in, Register) are at the top. The main content area has a blue header with the project name and a 'pip install cannlytics' button. Below this, a tagline reads 'Cannlytics = cannabis + analytics. Data pipelines, user interfaces, and the best statistics in the game. Made with ❤️!'. The left sidebar contains navigation links: 'Project description' (selected), 'Release history', 'Download files', 'Project links' (with 'Homepage'), and 'Statistics' (showing 11 stars, 6 forks, and 9 open issues/PRs). The main content area features a 'Project description' section with a colorful illustration of a planet, a tablet with a flame, and the text 'Cannlytics Engine'. Below the illustration, it says 'Simple, easy, cannabis analytics.' and provides the URL 'https://cannlytics.com'. At the bottom, there's a license section showing 'MIT', 'pypi', 'v0.0.11', and 'downloads: 128/month'. A footer note states: 'Cannlytics is a set of useful tools to get (wrangle), curate, augment, analyze, archive, and market cannabis data.'

Phishing attack against PyPI users [Read Details](#)

Search projects

Help Sponsors Log in Register

cannlytics 0.0.11 ✓ Latest version

`pip install cannlytics`

Released: Aug 20, 2022

🔥 Cannlytics = cannabis + analytics. Data pipelines, user interfaces, and the best statistics in the game. Made with ❤️!

Navigation

- Project description**
- Release history
- Download files

Project links

- Homepage

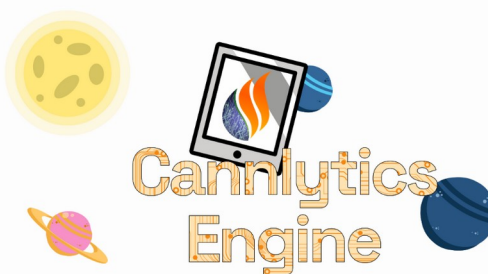
Statistics

GitHub statistics:

- ★ Stars: 11
- 🔗 Forks: 6
- 📄 Open issues/PRs: 9

View statistics for this project via [Libraries.io](#) or by using [our public dataset on Google BigQuery](#)

Project description



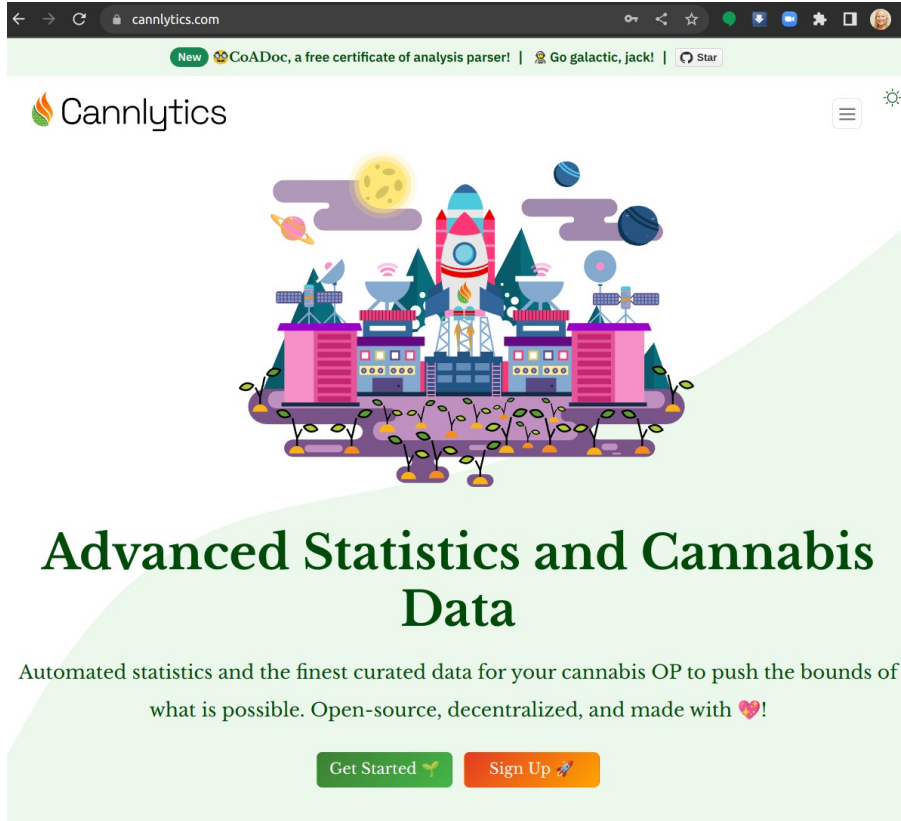
Simple, easy, cannabis analytics.

<https://cannlytics.com>

License: MIT | pypi | v0.0.11 | downloads: 128/month

🔥 Cannlytics is a set of useful tools to get (wrangle), curate, augment, analyze, archive, and market cannabis data.

Cannlytics engine CoADoc reads and parses COA Cannabis test lab “Certificates of Analysis” pdf file(s) and saves to .xlsx format.



The screenshot shows the Cannlytics homepage. At the top, a navigation bar includes a 'New' badge, a description of CoADoc as a free certificate of analysis parser, and links to 'Go galactic, jack!' and a 'Star' button. The main header features the Cannlytics logo and a menu icon. Below this is a large, colorful illustration of a cannabis farm with various equipment like satellites, a rocket, and a moon. The main heading reads 'Advanced Statistics and Cannabis Data'. A subheading states: 'Automated statistics and the finest curated data for your cannabis OP to push the bounds of what is possible. Open-source, decentralized, and made with ❤️!'. At the bottom, there are two buttons: 'Get Started' with a leaf icon and 'Sign Up' with a pencil icon.

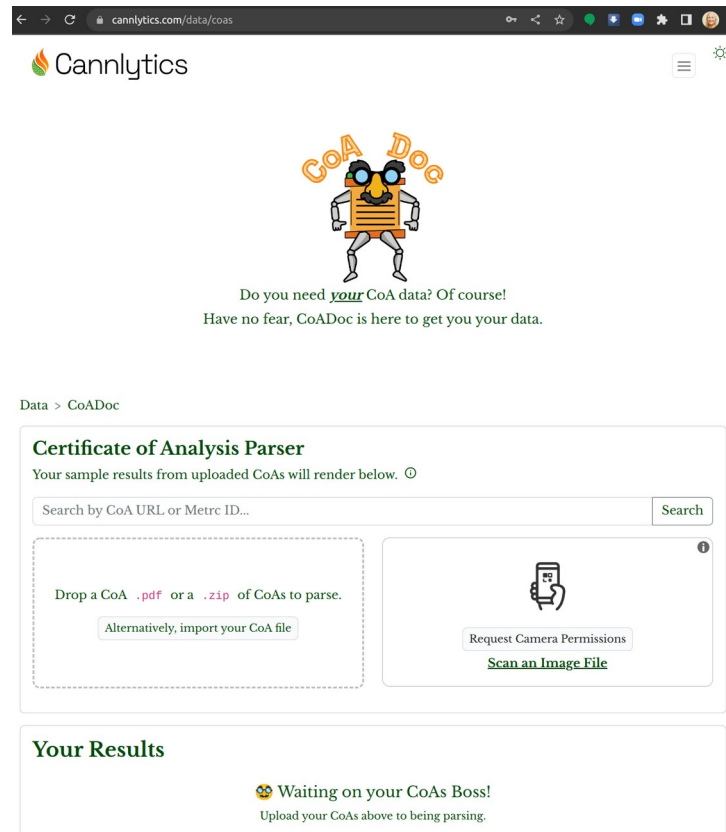
New 🚀 CoADoc, a free certificate of analysis parser! | Go galactic, jack! | Star

Cannlytics

Advanced Statistics and Cannabis Data

Automated statistics and the finest curated data for your cannabis OP to push the bounds of what is possible. Open-source, decentralized, and made with ❤️!

Get Started 🌿 Sign Up ✎



The screenshot shows the Cannlytics CoADoc interface. At the top, the Cannlytics logo is on the left, and a menu icon and settings icon are on the right. Below the logo is a cartoon robot character with 'CoA Doc' written above it. A message reads: 'Do you need your CoA data? Of course! Have no fear, CoADoc is here to get you your data.' Below this is a breadcrumb trail: 'Data > CoADoc'. The main section is titled 'Certificate of Analysis Parser' and includes a subtext: 'Your sample results from uploaded CoAs will render below. ⓘ'. There is a search bar with the placeholder 'Search by CoA URL or Metrc ID...' and a 'Search' button. Below the search bar are two main areas: a dashed box on the left with the text 'Drop a CoA .pdf or a .zip of CoAs to parse.' and a button 'Alternatively, import your CoA file'; and a box on the right with a camera icon, the text 'Request Camera Permissions', and a button 'Scan an Image File'. At the bottom, there is a section titled 'Your Results' with a message: '🤖 Waiting on your CoAs Boss! Upload your CoAs above to being parsing.'

Cannlytics

CoA Doc

Do you need your CoA data? Of course!
Have no fear, CoADoc is here to get you your data.

Data > CoADoc

Certificate of Analysis Parser

Your sample results from uploaded CoAs will render below. ⓘ

Search by CoA URL or Metrc ID... Search

Drop a CoA .pdf or a .zip of CoAs to parse.
Alternatively, import your CoA file

Request Camera Permissions
Scan an Image File

Your Results


🤖 Waiting on your CoAs Boss!
Upload your CoAs above to being parsing.

CoADoc read, parsed and placed our data into a table

canalytics.com/data/coas

Your Results

Download Results Clear



**Woodnote Farms:
Pistachio**
Flower, Inhalable
HUMBOLDT
SUNGROWERS GUILD,
LLC
02/19/2022

Custom CoA Parsing


At this time, CoADoc can only parse certificates of analysis (CoAs) from labs and LIMS with validated parsing algorithms. We've validated:

Labs	LIMS
Anresco Laboratories	Confident Cannabis
Cannalysis	TagLeaf LIMS
Green Leaf Lab	
MCR Labs	
SC Labs	
Sonoma Lab Works	
Veda Scientific	

If you want your favorite lab or LIMS added, then please email dev@canalytics.com and chances are that they can be included. Alternatively, because

canalytics.com/data/coas

Sample Results | Woodnote Farms: Pistachio



Woodnote Farms: Pistachio
Flower, Inhalable
HUMBOLDT SUNGROWERS GUILD, LLC
02/19/2022

Analysis	Compound	Result	Units	Status
terpenes	Nerolidol	0.1103	percent	
terpenes	Linalool	0.066	percent	
terpenes	Terpineol	0.0339	percent	
terpenes	Fenchol	0.0322	percent	
terpenes	trans- β -Farnesene	0.024	percent	
terpenes	Caryophyllene Oxide	0.0213	percent	
terpenes	β -Ocimene	0.0174	percent	
terpenes	Camphene	0.0166	percent	
terpenes	Borneol	0.0132	percent	
terpenes	Valencene	0.01	percent	
terpenes	Terpinolene	0.0061	percent	
terpenes	Sabinene Hydrate	0.0054	percent	
cannabinoids	THCa	26.875	percent	
cannabinoids	D9-THC	1.15	percent	
cannabinoids	CBGa	0.96	percent	
cannabinoids	CBCa	0.43	percent	
cannabinoids	THCVa	0.13	percent	
cannabinoids	CBDa	0.126	percent	
cannabinoids	CBG	<LOQ	percent	
cannabinoids	CBC	<LOQ	percent	
cannabinoids	D8-THC	ND	percent	
cannabinoids	THCV	ND	percent	
cannabinoids	CBD	ND	percent	
cannabinoids	CBDV	ND	percent	
pesticides	Cyfluthrin	ND	$\mu\text{g/g}$	PASS
pesticides	Cypermethrin	ND	$\mu\text{g/g}$	PASS
pesticides	Diazinon	ND	$\mu\text{g/g}$	PASS
pesticides	Dimethomorph	ND	$\mu\text{g/g}$	PASS
pesticides	Etiozazole	ND	$\mu\text{g/g}$	PASS
pesticides	Fenhexamid	ND	$\mu\text{g/g}$	PASS
pesticides	Fenpyroximate	ND	$\mu\text{g/g}$	PASS
pesticides	Flonicamid	ND	$\mu\text{g/g}$	PASS
pesticides	Fludioxonil	ND	$\mu\text{g/g}$	PASS
pesticides	Hexythiazox	ND	$\mu\text{g/g}$	PASS
pesticides	Imidacloprid	ND	$\mu\text{g/g}$	PASS
pesticides	Kresoxim-methyl	ND	$\mu\text{g/g}$	PASS

Download - saved our CoA data to an .xlsx worksheet

coa-data-2022-08-26-17-50-48.xlsx

	EB	EC	ED	EE	EF	EG	EH	EI	EJ	EK	EL	EM	EN	EO
1	acequinoc	acetamipr	azoxystro	bifenazate	bifenthrin	boscalid	captan	carbaryl	chlorantra	clofentezir	water_act	alpha_hun	piperonyl	dichlorvos
2	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	1E-09	0.175	1E-09	1E-09

How did Cannlytics CoaDoc do it?

- Lets step thru the code...

<https://github.com/cannlytics/>

Copyright (c) 2020-2022 Cannlytics and the Cannabis Data Science Team

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Please cite the following if you use the code examples in your research:

```
@misc{cannlytics2022,  
  title={Cannabis Data Science},  
  author={Skeate, Keegan and Rice, Charles and O'Sullivan-Sutherland, Candace},  
  journal={https://github.com/cannlytics/cannabis-data-science},  
  year={2022}  
}
```



Cannlytics

Simple, easy, end-to-end cannabis analytics. Open source. Made by scientists for scientists.

<https://cannlytics.com> @cannlytics contact@cannlytics.com

Unfollow

[Overview](#)
[Repositories](#) 20
 [Projects](#) 3
 [Packages](#)
[Teams](#)
[People](#) 2

readme.md

Cannlytics

Advanced Statistics and Cannabis Data

Introduction

Cannlytics is a suite of free software for the cannabis industry.

Contributing

See the [main Cannlytics repository](#) to begin.

View as: Public ▼

You are viewing the README and pinned repositories
as a public user.

People



Top languages

● Python ● JavaScript ● Dart ● CSS

● HTML

Most used topics

Manage

cannabis cannabis-app cannabis-data
cannabisapp python

Pinned



🔥 Cannlytics = cannabis + analytics. Data pipelines, user interfaces, and the best statistics in the game. Made with ❤️

Python ☆ 11 🍷 6



cannabis-data-science Public

🚀 Cannabis Data Science repository powered by 🔥 Cannlytics. 🤝 Meetup, code, and advance cannabis science 🌿. Join the fun!

Python 11 3