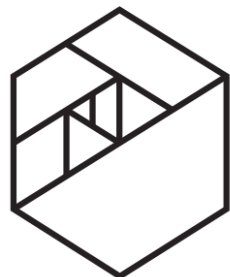# Data Science!

METIS

datascope

Why are we here?

Goal. To enrich and broaden your skills and confidence to ask & answer questions using data, statistical models, code.

Why are we here?

Goal. To enrich and broaden your skills and confidence to ask & answer questions using data, statistical models, code.

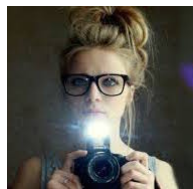**We are data scientists.**
This is data science.

What is a data scientist?

A data analyst who lives in California

Someone better at statistics than any software developer, better at software development than any statistician
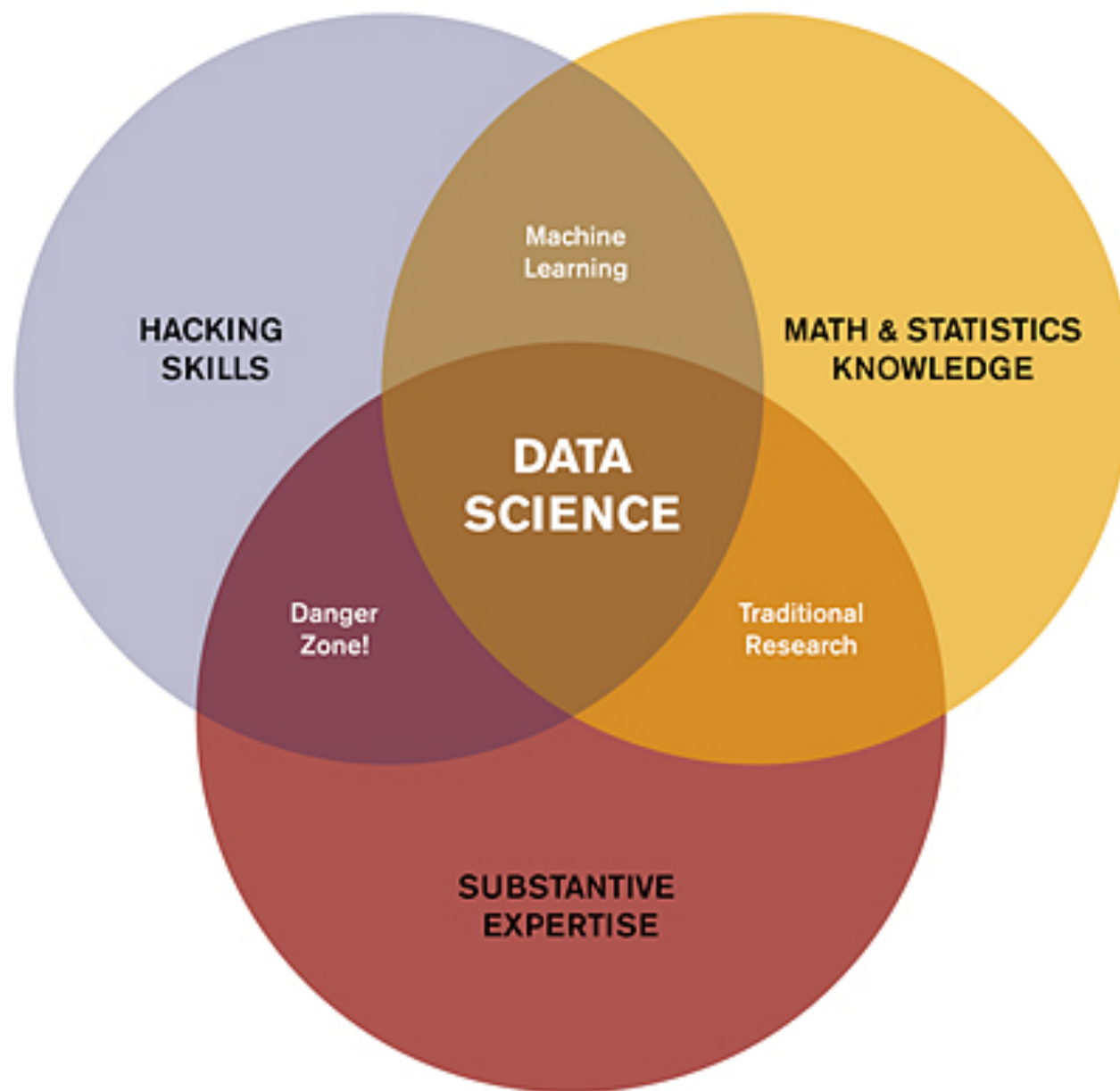
A business analyst who lives in New York

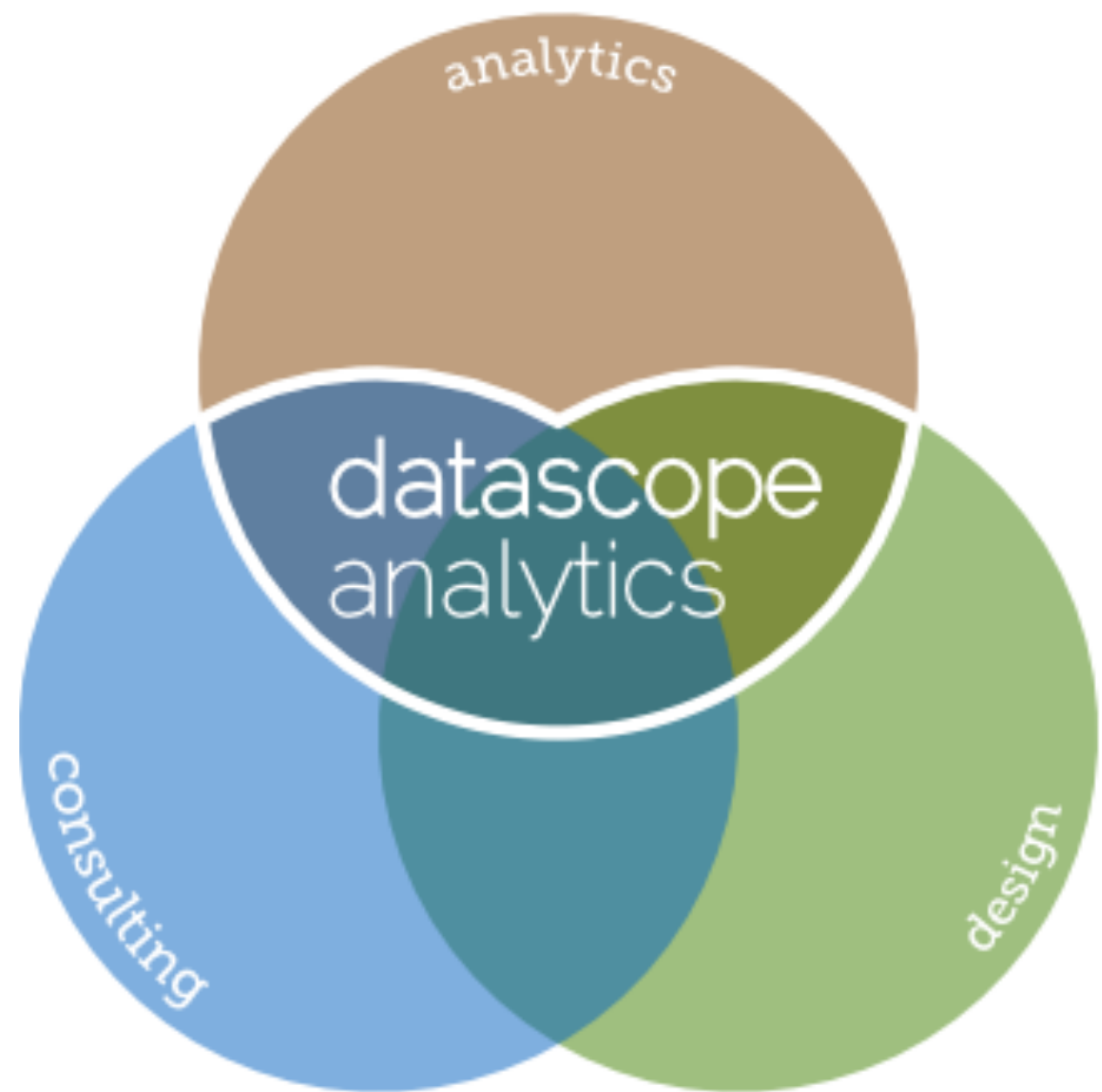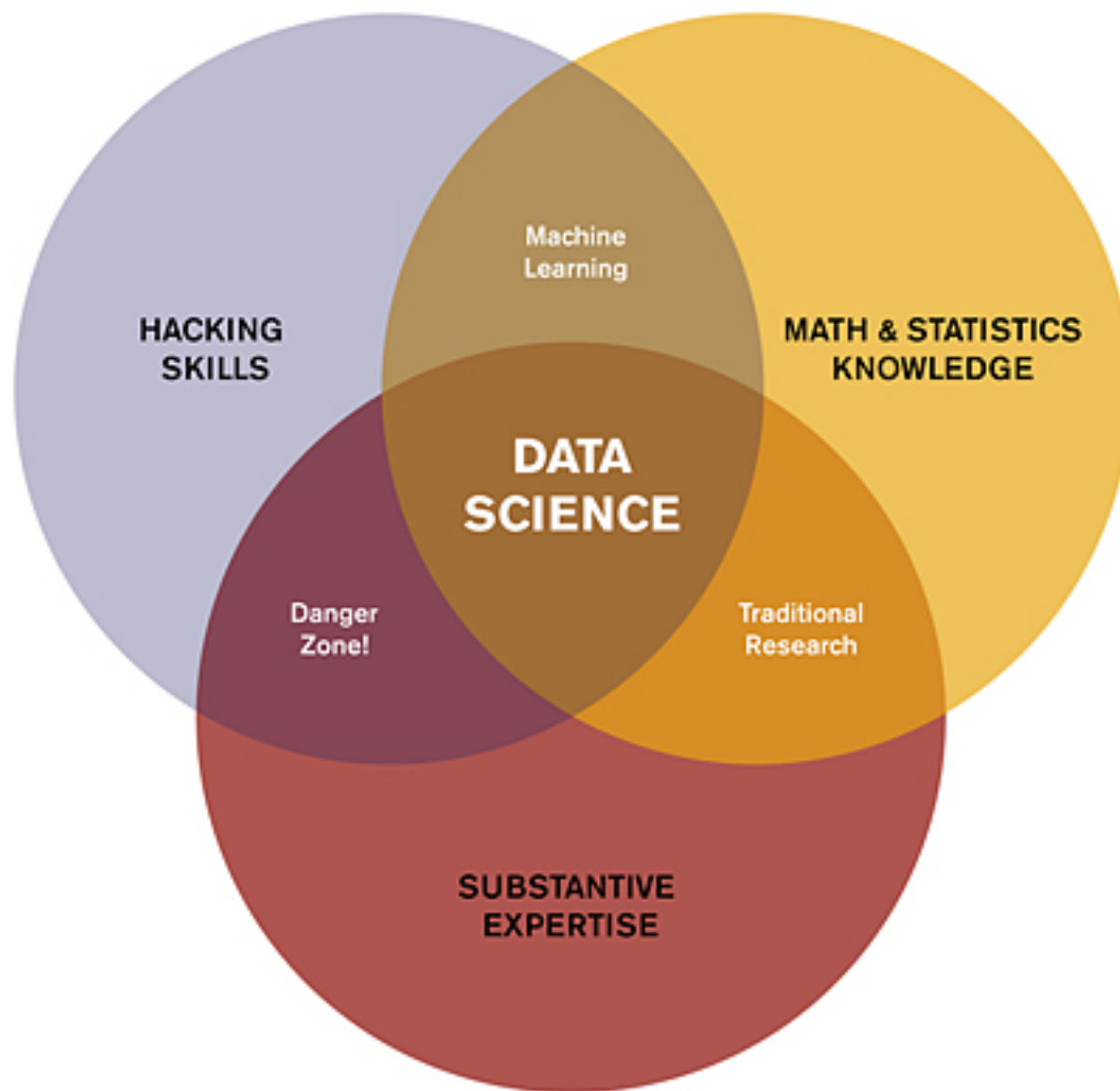A statistician who lives in San Francisco

Performer of statistics on a Mac

What is a data scientist?

What is a data scientist?

Skill Domains

Statistics /
Machine Learning

Programming

Communication

Design

Not skills

Curiosity

Creativity

The Anatomy of a Data Science Project

Domain	Design	Data	Algorithms	Tools	Visualize Communicate

Domain

Finding the relevant questions

The right metrics

Knowing what's been done

Design

Start with a question, not data

Iterative design process

Moving targets

Design

Start with a question, not data

Iterative design process

Moving targets

Data

Find Data

Clean Data

Manage Data

Data

Find Data

Clean Data

Manage Data

BIG DATA

Machine Learning

Statistics

Applied Math

Algorithms

Open Source Tools

Python
Pandas, Scikit.learn
SQL, Mongo
Javascript, d3, Flask
Hadoop, Spark

Tools

Google

Tools

Easy to read graphs

Interactive dashboards

Explaining well, adapting to the audience

Visualize
Communicate

Easy to read graphs

Interactive dashboards

Explaining well, adapting to the audience

Visualize
Communicate

Ultimate Goal & Product: Insights

Domain Design Data Algorithms Tools **Visualize** Communicate
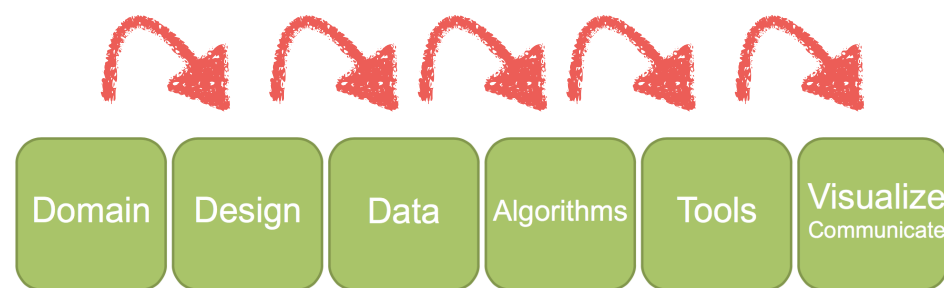
A Project-based course

Week 1

Domain → Design → Data → Algorithms → Tools → Visualize Communicate

Week 2

Domain → Design → Data → Algorithms → Tools → Visualize Communicate
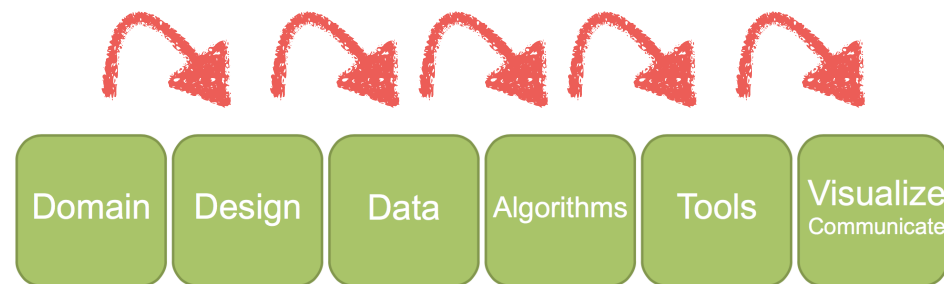
Repeat the whole process twice

Learn by actually doing it

Active learning is super effective

Doesn't work without real effort

Format

🕘 9:30am -12:00pm

🕐 1:30pm - 5:00pm

Follow the white rabbit.

Lessons, pair exercises, challenges, **project work**

Zoom in on the Format

 Focused time

 Peer and community support

 Self-learning with a safety net

Focused Time

Come to class every day

Work on assignments

Go for the best project outcomes you can

Be curious and creative

 # Community Support

Nobody knows everything

We don't know everything

Different experience levels in different areas

Learn by teaching: Help each other.

Self-learn with a safety net

Nobody knows everything

Lessons/knowledge important

The skill to walk through when you hit walls is much more important: **the hacker ethos**.

Self-learn with a safety net

Lessons can't cover everything

Ability to quickly learn is crucial

Best way to get it:
Build the foundation.
Keep getting stuck.
Practice & repetition.

Self-learn with a safety net

Stuck. Ask peers. Ask google. Learn. Proceed.

Safety net: Instructors Ask us.

Agency:
Us ——► You

Zoom in on the Format

Focused time

Peer and community support

Self-learning with a safety net

Why are we here?

Goal. To enrich and broaden your skills and confidence to ask & answer questions using data, statistical models, code.

We are data scientists.
This is data science.

You.

Who are you?

What do you do?

What are your expectations?