A practitioner's view on data science tools

Actuarial Data Science Après-Midi October 7, 2020

David Lüthi

Goal

- Know your opportunity set

- Ask the right questions

- Understand your peer's setup

Intro

Speaker   David Lüthi

Head Personal Lines Pricing P&C, Zürich Schweiz

linkedin.com/in/davidluethi

Disclaimer   My personal view

No claim to completeness

Scope   In scope: Data Science tools

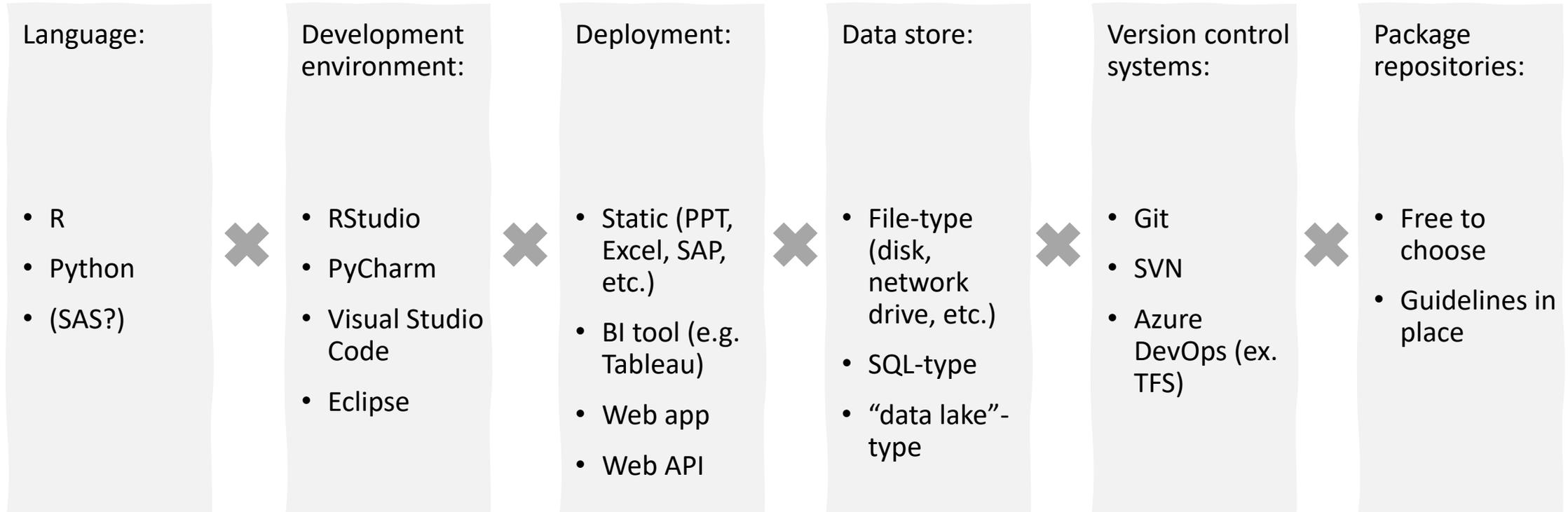Out of scope: Data Science studios & compute infrastructure

The actuarial data scientist's mandate

- Which language?
- Which development environment?

- Where/how to deploy?
- Slides or web APIs?

"[…] transform data into insights with your team."

- Which data store?
- What computing infrastructure?

- How to share code & collaborate?
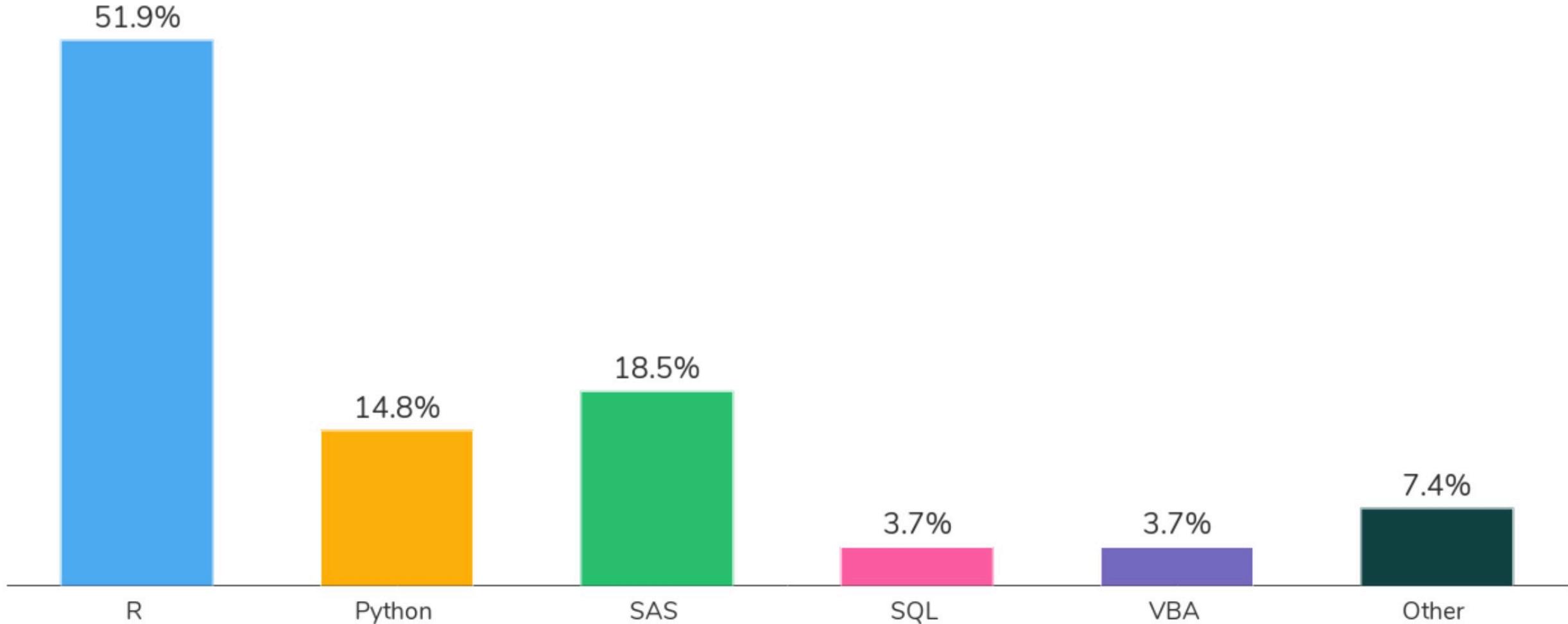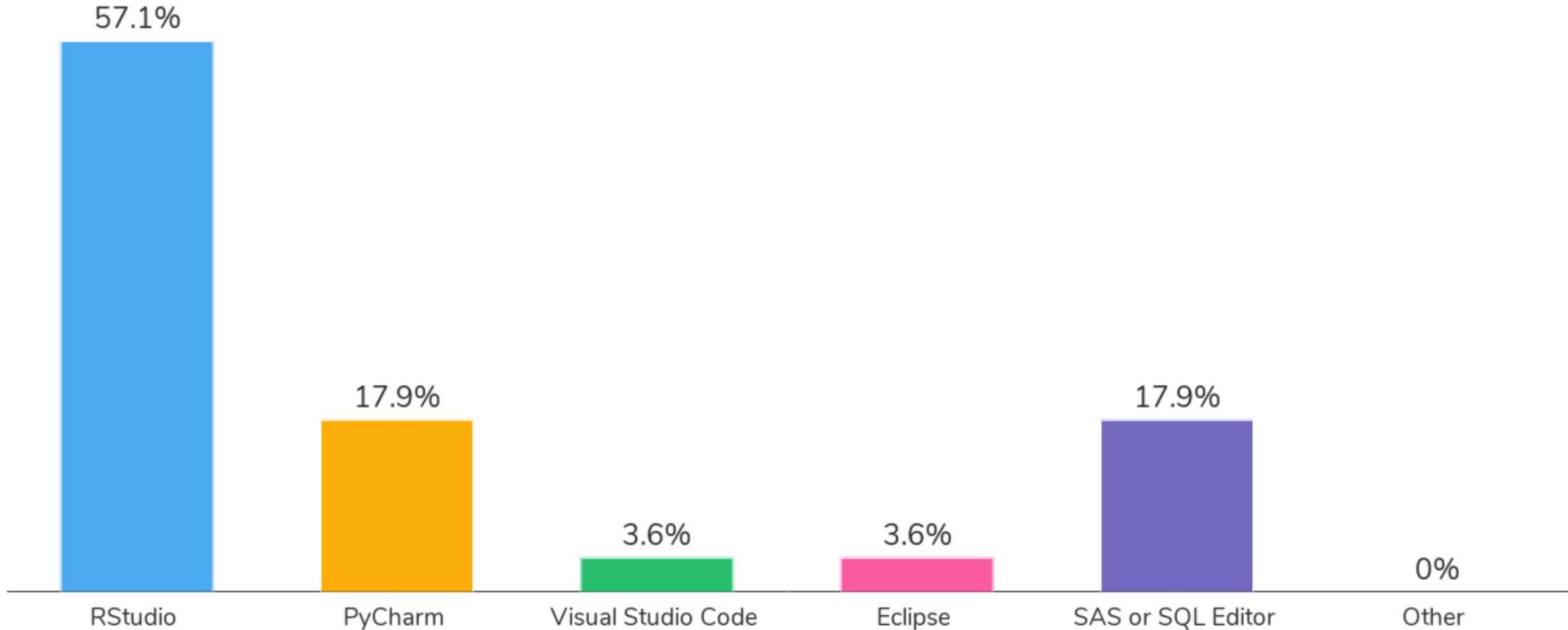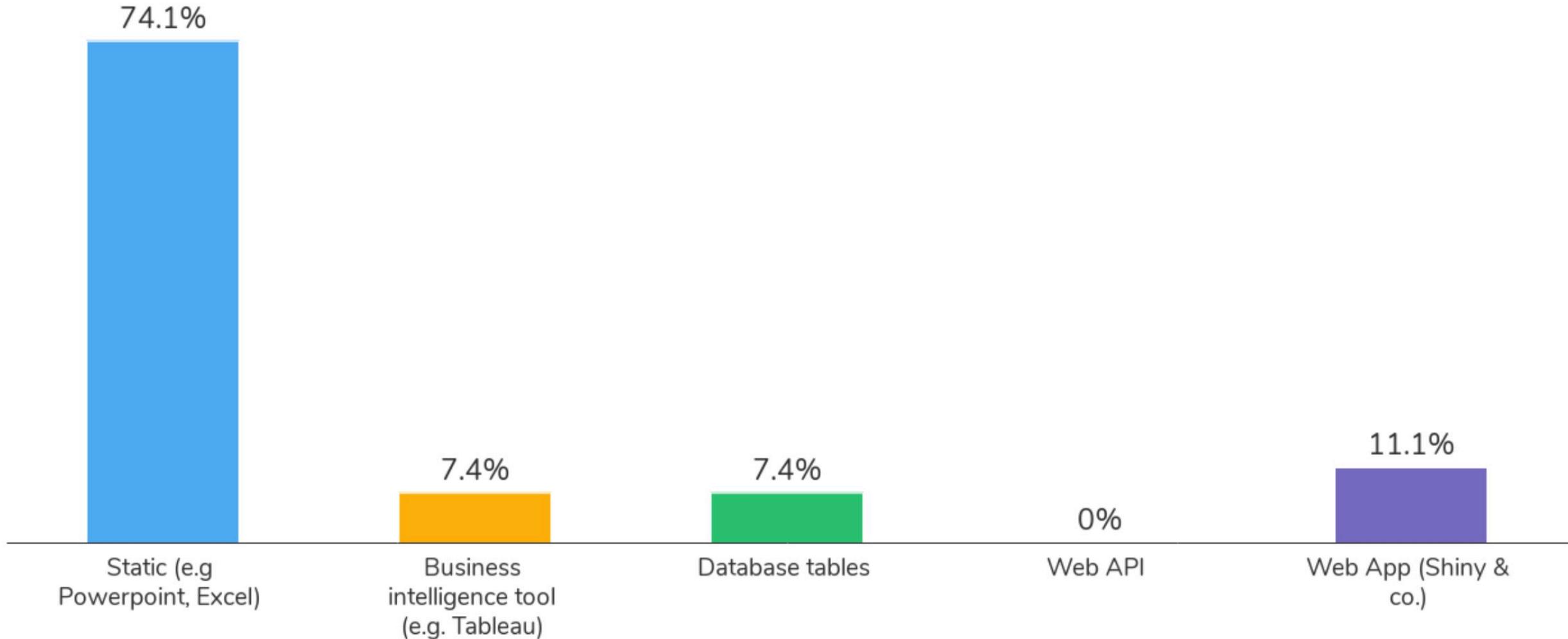- What version control system?

A opportunity subset

Language:

- R
- Python
- (SAS?)

Development environment:

- RStudio
- PyCharm
- Visual Studio Code
- Eclipse

Deployment:

- Static (PPT, Excel, SAP, etc.)
- BI tool (e.g. Tableau)
- Web app
- Web API

Data store:

- File-type (disk, network drive, etc.)
- SQL-type
- "data lake"-type

Version control systems:

- Git
- SVN
- Azure DevOps (ex. TFS)

Package repositories:

- Free to choose
- Guidelines in place

What's your setup?

ahaslides.com/ADSAM

What's your primary language?

R	Python	SAS	SQL	VBA	Other
51.9%	14.8%	18.5%	3.7%	3.7%	7.4%

Which development environment are you using?

57.1%	17.9%	3.6%	3.6%	17.9%	0%
RStudio	PyCharm	Visual Studio Code	Eclipse	SAS or SQL Editor	Other

28

Where do you mainly deploy your results?

27

What kind of data store do you use predominantly?

67.9%

25%

7.1%

SQL-type

File-type (local disk, network drive, etc.)

"Data lake"-type

28
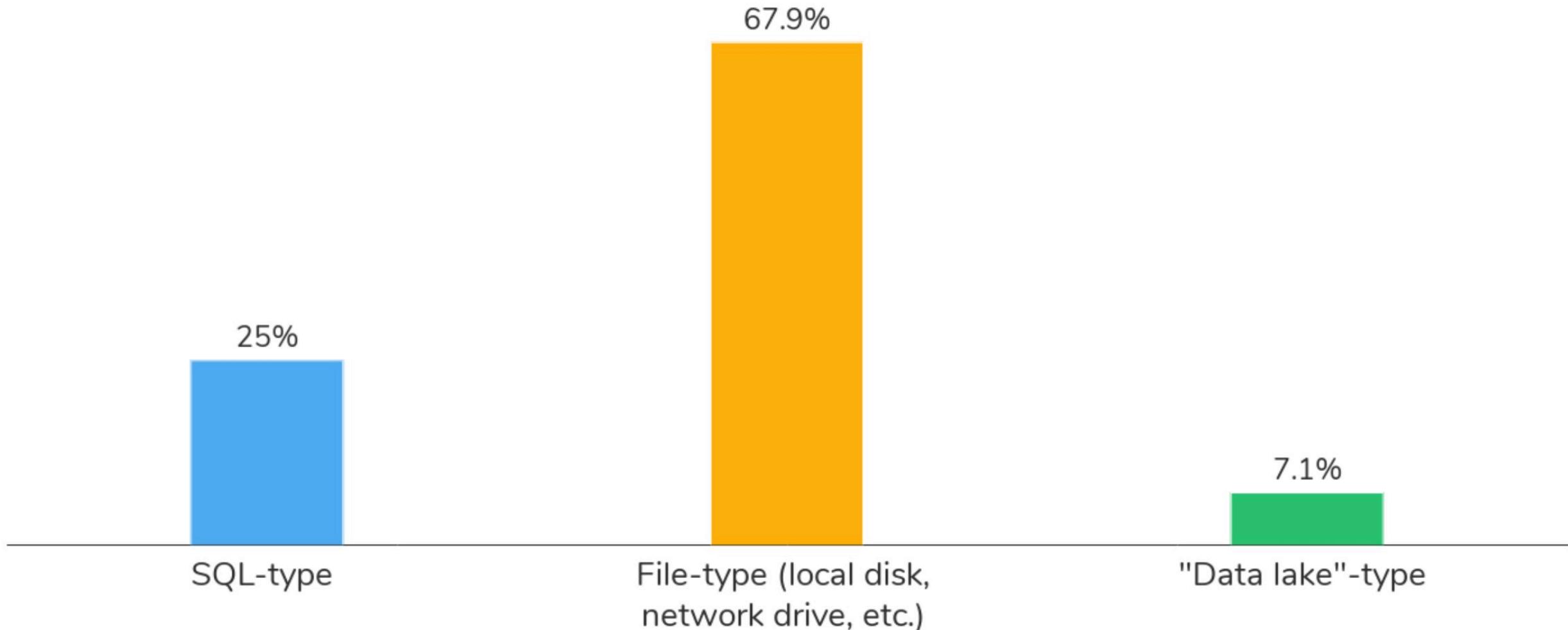
What kind of computing infrastructure setup do you use?
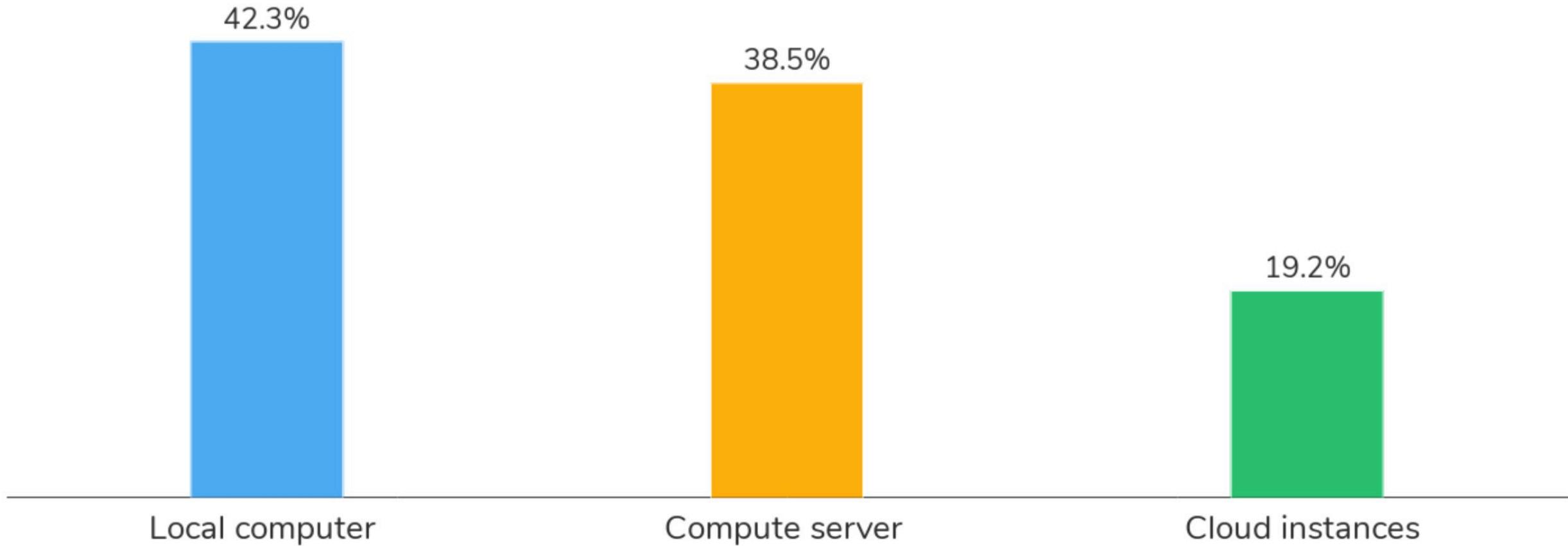
Which version control system are you using?

	Git	SVN	Azure DevOps	None
	33.3%	7.4%	0%	59.3%
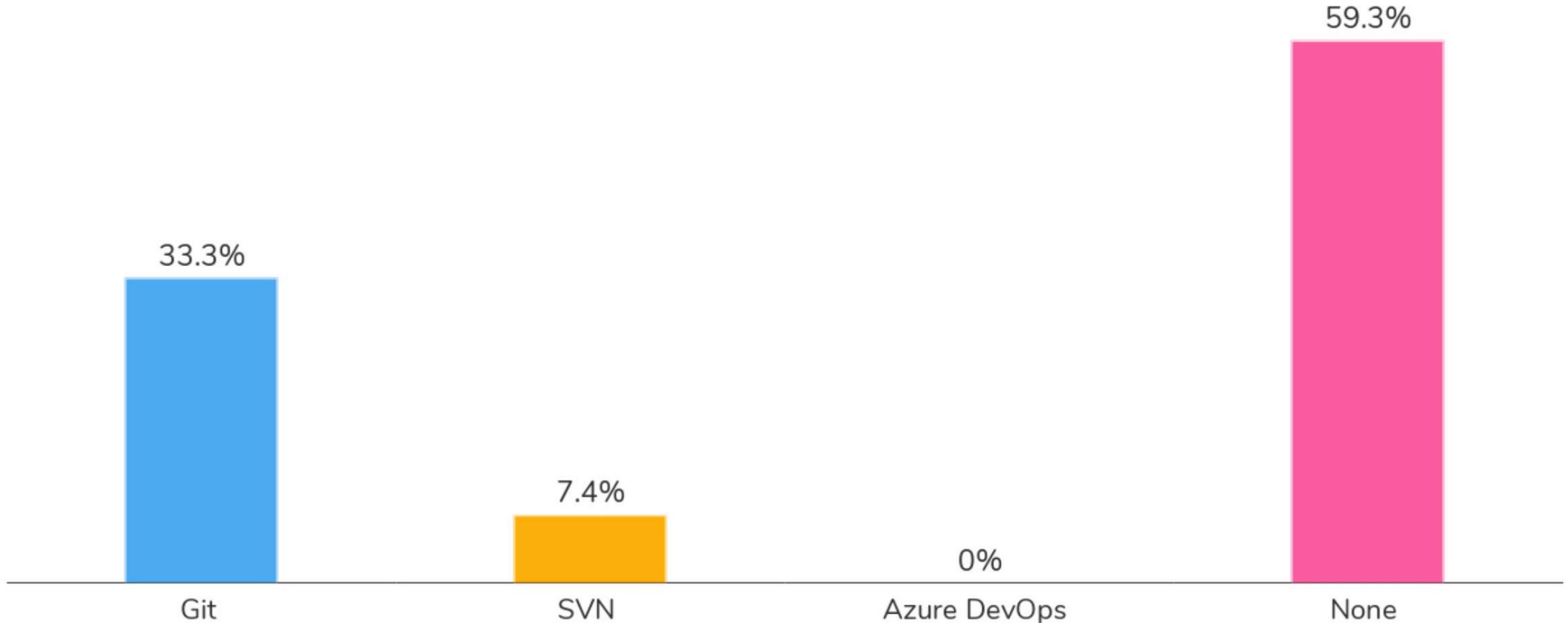
Five key questions

1) What's the Data Science literacy of your team?

2) How important is collaboration?

3) What types of business problems do you solve?

4) What's the level of support from IT/Ops?

5) How computationally intensive are your problems?

My professional experience

…some stories…

Thank you!