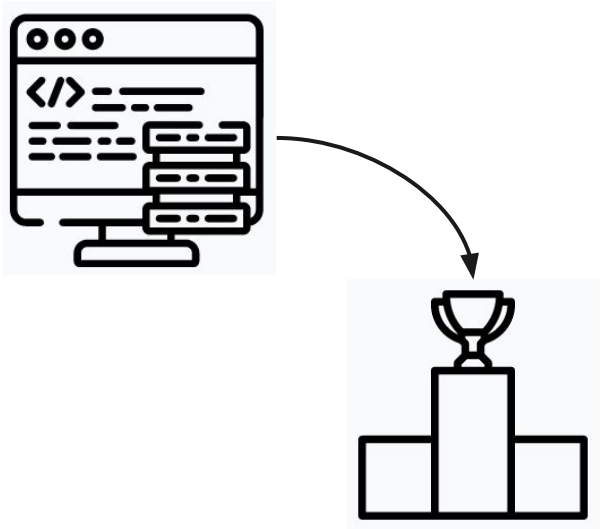
Welcome to the insurance pricing game

by Ali Farzanehfar and Florimond Houssiau

April 8th, 2019.

The pricing game is a data science competition

1. We are from the Data Science Institute
2. We will provide PIZZA today (ORDER NOW PLEASE)
3. Prize (provided by ICDSS): up to £100 Amazon voucher
4. A link to the slides will be emailed to you
5. This is being live streamed



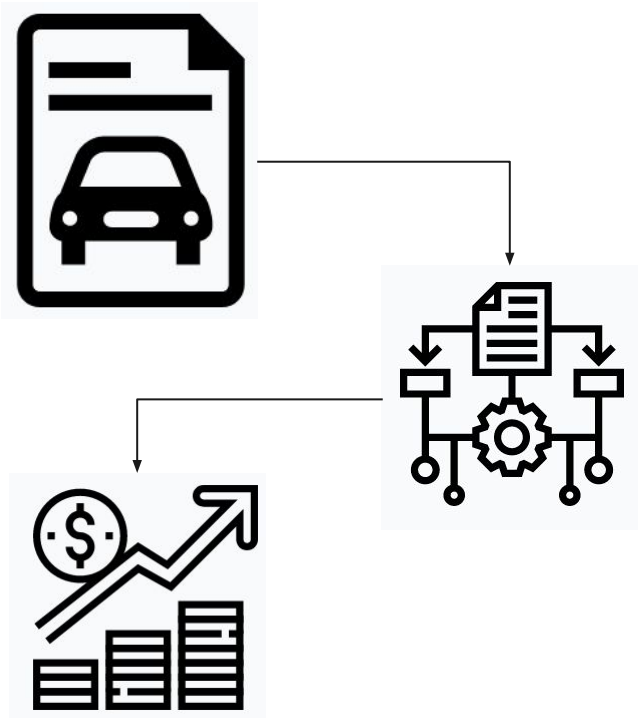
The story:
You are an insurance company

You have been hired by a new car insurance company

1. Your company has historical data on insurance
    a. Contract info + claims info

2. You must come up with a way to algorithmically generate premium quotes —> Insurance pricing

3. Success is measured on unseen data



Customers will always pick the cheapest premium

Customers ↓

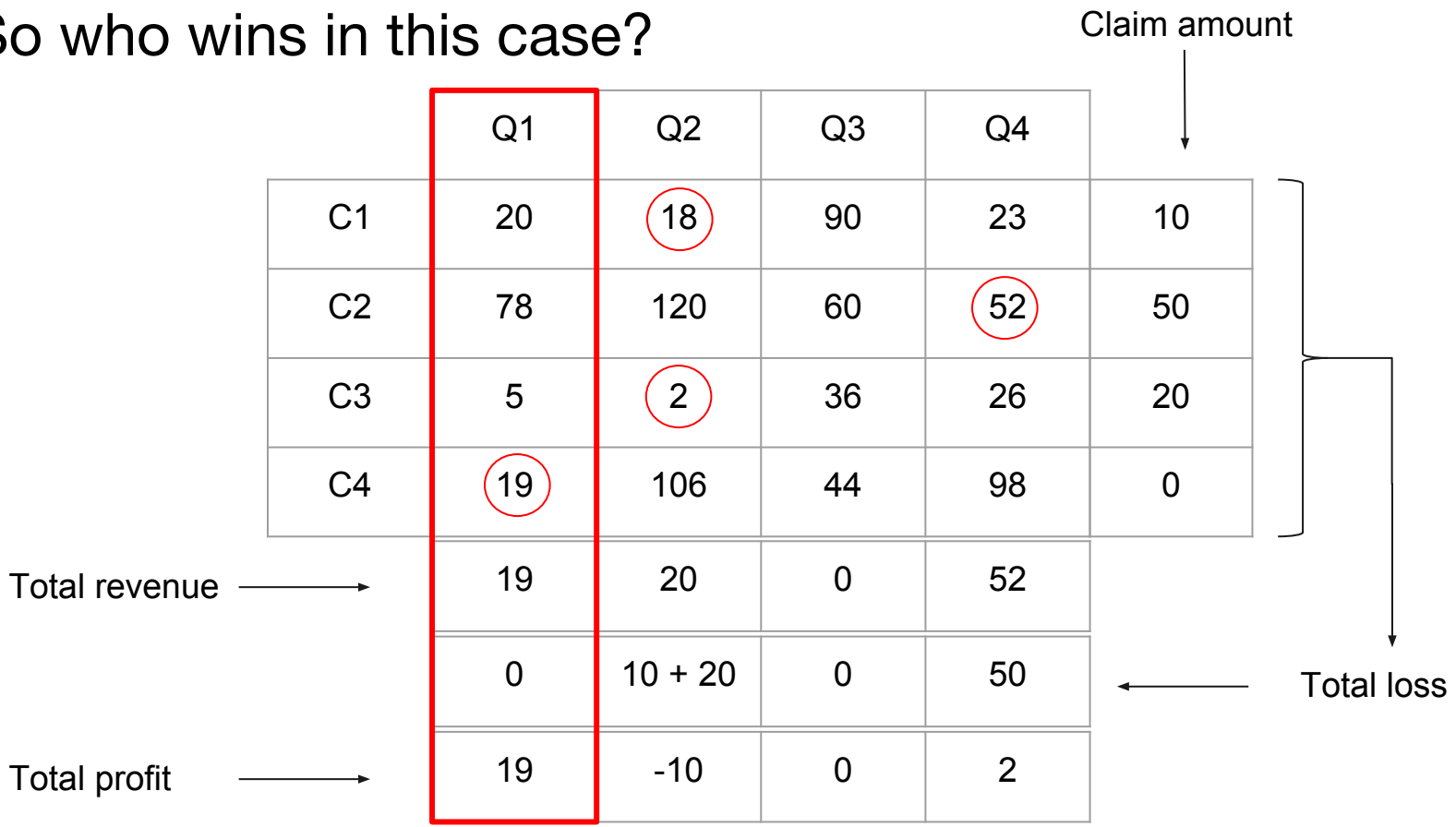
	Q1	Q2	Q3	Q4
C1	20	18	90	23
C2	78	120	60	52
C3	5	2	36	26
C4	19	106	44	98
Total revenue →	19	20	0	52

← Pricing models

Some customers will have accidents and file claims

	Q1	Q2	Q3	Q4	Claim amount
C1	20	18	90	23	10
C2	78	120	60	52	50
C3	5	2	36	26	20
C4	19	106	44	98	0

So who wins in this case?



The rules of the game: Sharing, deadline, versions

1. Your will be given data that you cannot share with others
    a. Your data is your company's competitive advantage
2. You are to submit a model for pricing by **the end of today's session**
    a. We will meet again next week to distributed the prize
    b. We MIGHT have a second stage next week depending on participation
3. You are asked to code in **Python 3.5+** (with a restricted set of libraries)
    a. Set of supported libraries available here: github.com/alfarzan/pricing_game/

The rules of the game: profit

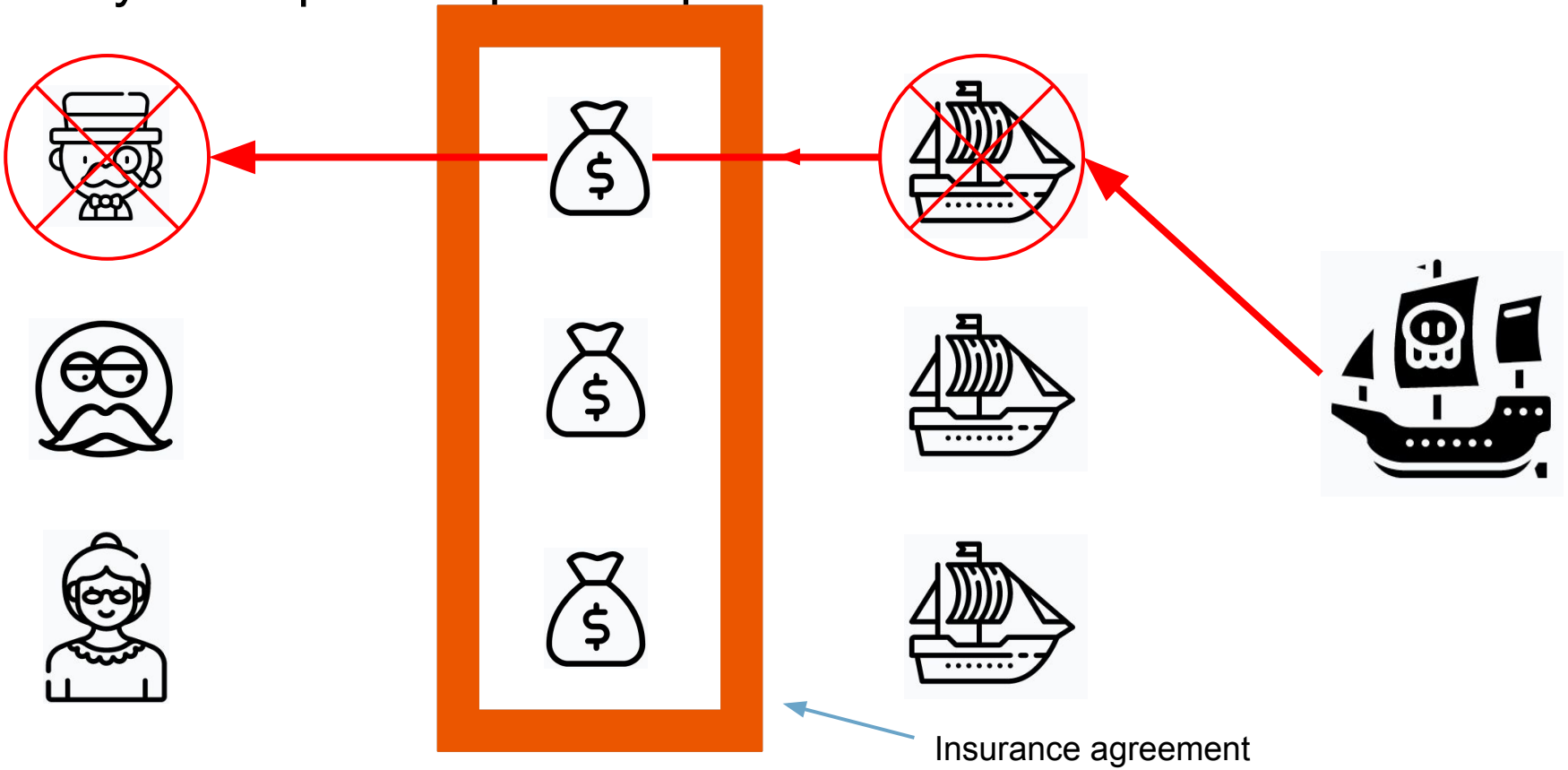
1. Your model cannot lose money on your training data
    a. If this is not satisfied, you cannot submit

2. Your model must make at least 10% profit on your training data
    a. If this is not satisfied but (1) is satisfied, we will linearly scale all your premium values to ensure 10% profitability

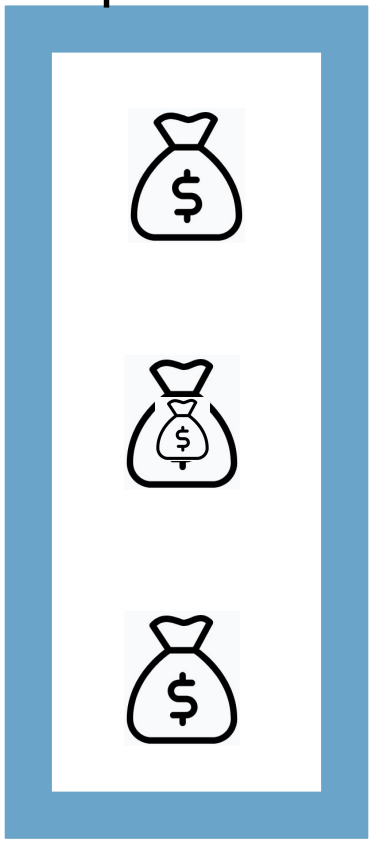
1st Place Prize:
£100 Amazon voucher
(if you submit a first model today)

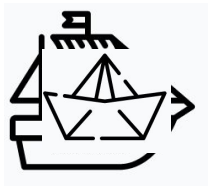
Insurance pricing: a primer

Toy example: Ships and pirates



Toy example: Ships and pirates





Insurance pricing is about making money from risk

1. Typical actuarial data set has:
    a. Contracts —> info on people signing up for insurance
    b. Claims —> how much they cost the insurance company at the end

2. Aim is to use details in the contract to predict two things:
    a. Probability that an accident will occur —> probability of making a claim
    b. How much will that accident cost —> amount of a claim

Pricing models: predicting risk and making profit (1)

1. A given pricing model Q has two components:

$$Q = \mathcal{M} + P$$

Underwriting: This is your best estimate of the total amount of claims in your training data. Your total cost.

Profit making: This is how you choose to adjust your underwriting prices to ensure you win in a competitive environment.

Goal:

Maximise profit in the market

A word of caution: You are in competition in a market

- You only win a customer if your price is **cheaper than everyone else** for that customer

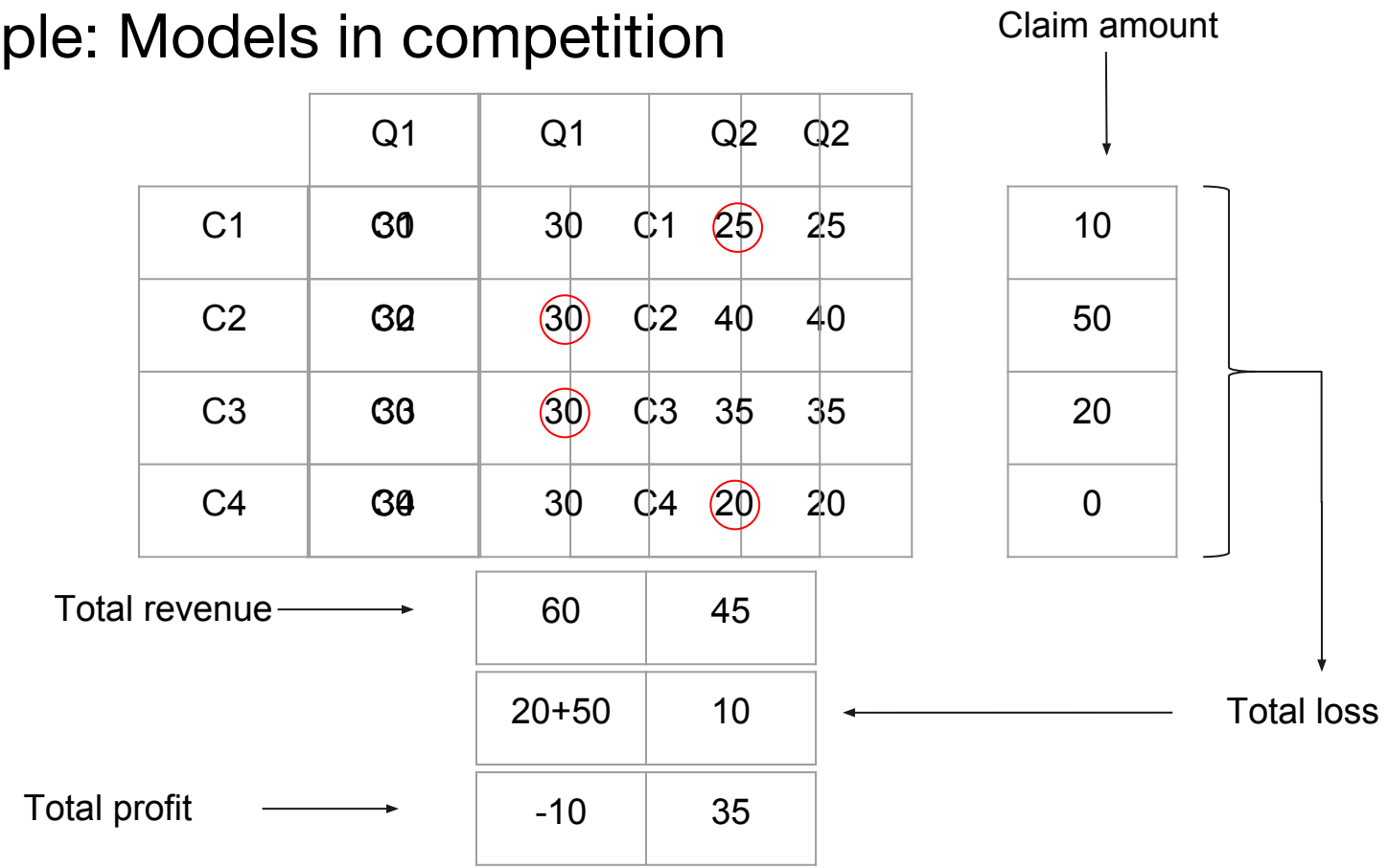
Example: Models outside of competition

	Q1
C1	30
C2	30
C3	30
C4	30
Total revenue	120
	80
Total profit	40

	Q2
C1	25
C2	40
C3	35
C4	20
	120
	80
	40

Claim amount
10
50
20
0
Total loss

Example: Models in competition



Technical details

The data set: Liability motor insurance

1. Real motor insurance data from 2007 in France
2. All data is from the same year
3. This is **liability insurance data** —> Claim amount specifies damage to third party
    a. If Alice registers a claim of £5, this means she had an accident and caused a damage worth £5 to **another car**
4. All customers are **legally required** to buy a policy (no customer can opt out)
5. Data dictionary available on: github.com/alfarzan/pricing_game/

Inginious set up (Live demo 10 min)

1. We will all sign up now on Inginious:
   pricing-game.dsi.ic.ac.uk/
   a. Course password: `PricingGame43`

2. If you're a team, only one member can sign up

Download the testing suite

1. Download the testing suite repository: github.com/alfarzan/pricing_game/
2. Make sure you are running Python 3.5+
3. In the repository you will find:
    a. `model.py` —> a template file for you to fill in with your model code
    b. `test_your_model.py` —> a file that tests your models' profitability
    c. `package_support.md` —> A list of supported packages
    d. `data_dictionary.md` —> The data dictionary

You are required to submit 3 files

1. `model.py` —> your model code
2. `model.pickle` —> your trained model (produced by `model.py`)
3. `description.txt` —> a description of your model

Obtaining your company data set

1. A dropbox link has been sent to your team email
2. Your company’s data set is available in the linked folder
3. Download it now and share it with your team members

Using Inginious

Toy example: A simple linear model

1. Let's look in the template file `model.py`
    a. downloaded from github.com/alfarzan/pricing_game/
2. It includes a toy model
    a. Logistic classifier to predict who will have a claim
    b. Linear regression to predict by how much
3. Let's write this

Making the simple model: What I need (live demo)

1. `model.py` —> your model code
2. `model.pickle` —> your trained model (produced by `model.py`)
3. `description.txt` —> a description of your model

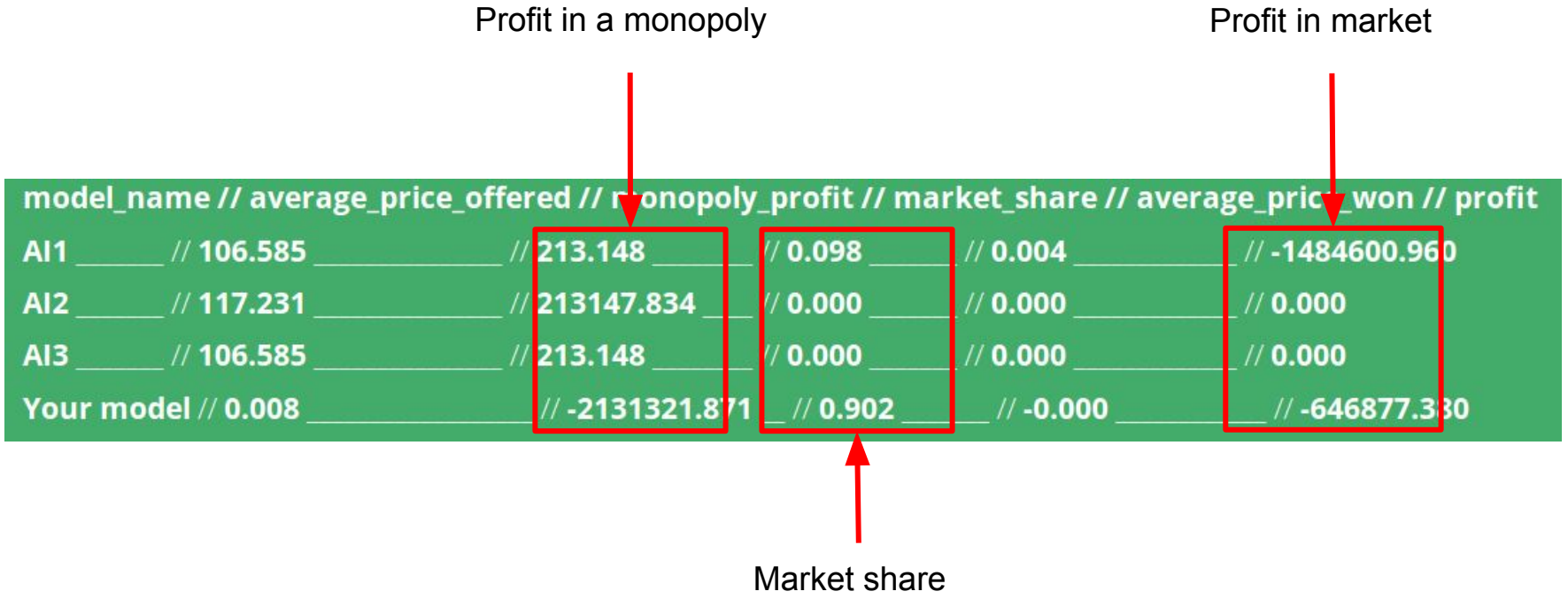
Testing the simple model: `test_your_model.py`

Upload to inginious (live demo)

1. Profit checker on Inginious

2. AI market on Inginious

AI market: the effect of competition



Final details

About your pricing models

1. You are free to **use any model** you wish (CNN, GLM, tree-based, your local psychic, etc)

2. In this game there are **no ethical issues** with regards to your pricing strategy
    a. You could pick to offer very expensive premiums to men and cheap ones to women

Deadlines, Q&A and, submission

1. Final submission deadline is next Monday at 2359 (April 15 2019)

2. We will meet back here on April 17 at 1500 to reveal the winner and distribute the prize

You have until 2130 to submit your first model!

Good luck!