Roadmap: first 100 days

Madalina Ciortan

Overview

Day 1

Day 100

Understand status quo

Address blocking points and low-hanging fruits

Iterative improvements

Prepare roadmap for next 100 days

Step 1: understand the current state

- Inventory data sources and data consumers
- Overview of the technical architecture for data storage and transformation
- Overview of the operationalization of Machine Learning models (MLOps)
- Identify stakeholders
- Understand the roles and responsibilities of the involved teams
- Clarify the business objectives and the underlying pain points
- Staff skill set evaluation

Step 2: Address blocking points and low-hanging fruits

- Move from excel processing to a more scalable approach on a data platform
    - backend python logic
    - CI/CD
    - Evaluate robust data storage options (cloud, databases, datalake, etc.)
- Automate data ingestion, ETL, validation
- Provide training (if needed)
    - Usage of the data platform

Step 3: Agile iterative improvements

- Set-up short feedback loops from the business
- Iterative improvement of the forecasting algorithm
- Iterative improvement of the web interface
- Start customer segmentation use-case
- Extend the forecasts to the warm-water shrimp use-case

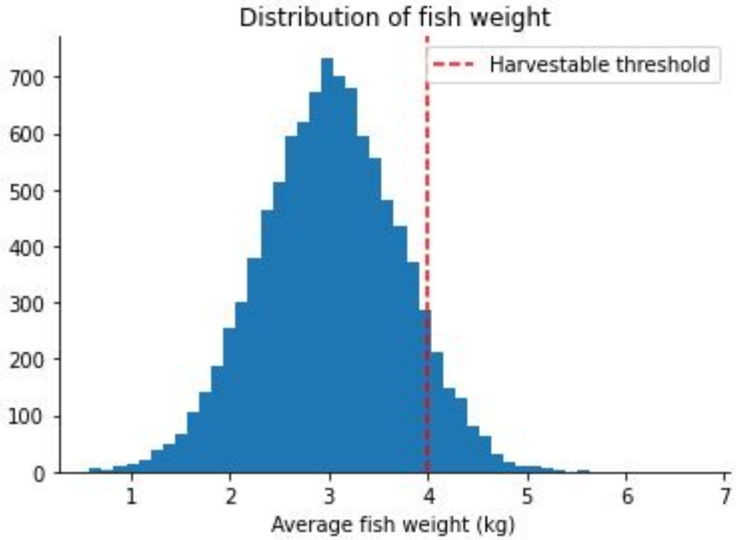
Step 4: New roadmap

- Define new long-term objectives
  - New projects
  - Improvements to the architecture
  - Training

Assignment 2.1

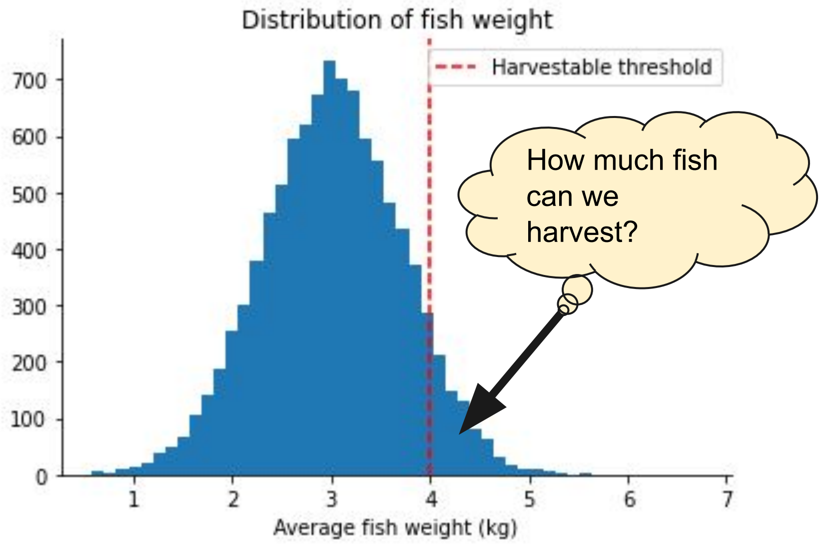
How much fish can we harvest at a given time?

- 1 Cage with average fish weight = 3kg
- We harvest all fish > 4kg



How much fish can we harvest at a given time?

- 1 Cage with average fish weight = 3kg
- We harvest all fish > 4kg



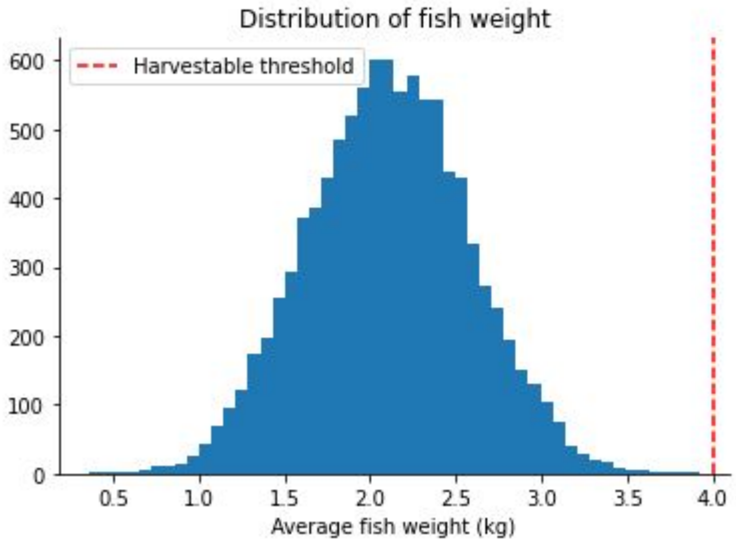
A specific scenario

Biomass = 281952

Average weight = 2.1

Standard deviation = 0.475

Harvest threshold = 4



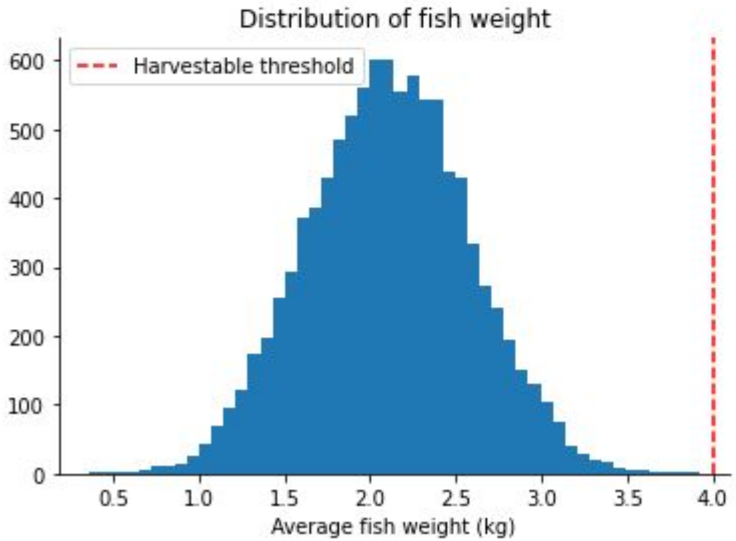
A specific scenario

We can compute how likely it is to find fish > threshold

=> Probability 0.003 %

=> Expected harvest = Biomass * Probability of harvest

We expect 8.9 kg to be harvested



Assignment 2.2

Simulate the harvest over 1 year

- Each month the fish grows by 11.2%
- We know only the total biomass and the number of individuals at the start of the first month
- All fish over 4kg is harvested at the end of the month

How much is harvested over the year?

Simulate the harvest over 1 year

- We know only the total biomass and the number of individuals at the start of the first month

  => We can compute the mean at the beginning of the month as biomass / nb individuals

  Mean = 317895/156077 = 2.03

  => We can compute the mean the end of the month, after growth

  Mean = Mean + Mean * 11.2% = 2.26

  => We can determine from the lookup table the std and compute using the previous approach the harvest = 48.8 kg

Simulate the harvest over 1 year

- The total biomass is reduced after harvest:

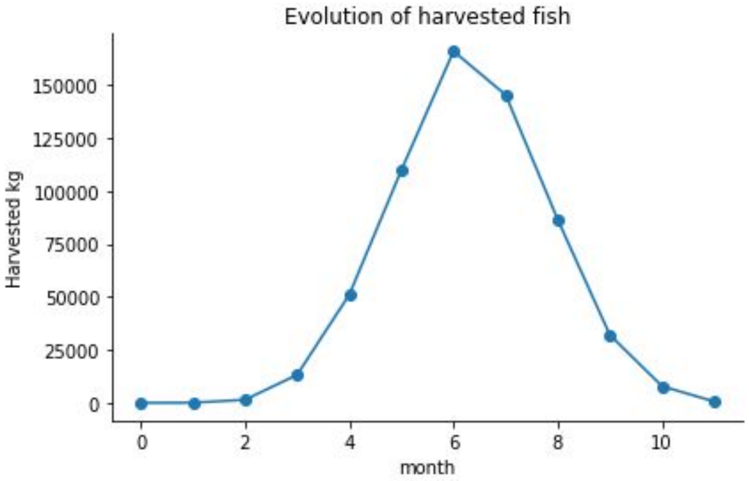
  New biomass = Old biomass - harvest = 353 450

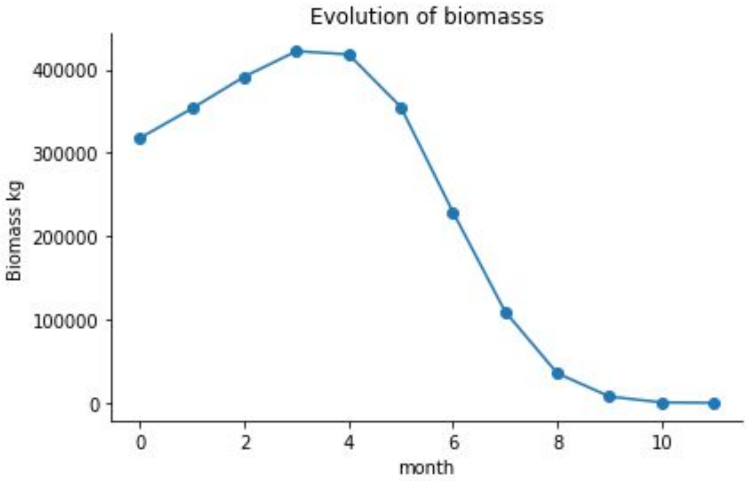
- We apply the same logic 12 times using the new biomass and the new mean

Results

	month	harvested	biomass	mean	total_harvest
0	0.0	0.000000	317895.000000	2.036783	0.000000
1	1.0	48.873730	353450.366270	2.264903	48.873730
2	2.0	1452.997345	391583.809948	2.518572	1501.871074
3	3.0	13235.828141	422205.368521	2.800652	14737.699216
4	4.0	51085.879097	418406.490698	3.114325	65823.578313
5	5.0	109898.700771	355369.316885	3.463129	175722.279084
6	6.0	165952.111148	229218.569228	3.851000	341674.390232
7	7.0	145271.607476	109619.441506	4.282312	486945.997708
8	8.0	86364.384199	35532.434756	4.761931	573310.381907
9	9.0	31949.225632	7562.841817	5.295267	605259.607538
10	10.0	7848.039548	561.840553	5.888337	613107.647086
11	11.0	624.766694	0.000000	6.547831	613732.413780

Evolution of harvest and biomass





Assumptions

Fish growth is constant 11%, not a function of age

Fish growth is not a function of available space in the cage.

If fish growth is constant 11%, the optimal strategy might be to harvest all fish in december when they weigh between 4 and 14kg

No fish dies in cage.

Other questions

- Can we propose strategies to optimize the harvested quantity?
- Knowing the expected supply, can we provide dynamic pricing insights and recommendations ?

Questions?