# Fresh grads - Software Engineer Assignment

Thank you for answering these questions to the best of your competences. The purpose of this assessment is to evaluate the things you know as well as your analytical thought process. Please return your completed assignment preferably within 7 days of reception (maximum 10 days). Depending on your proficiency level, answering all questions should take you between 6 and 12 hours of work time. The questions are broad, so prioritize your efforts aggressively on key information and simple solutions. If a question is unclear, feel free to rephrase it or lay new assumptions, but don't forget to document them. It is advised to read and understand all questions before starting to answer them.

For more information on the Nexthink product visit our [website](), [blog](), and [documentation](). Before starting the assignment, watch these short videos created by our partners, in order to understand the main components of our product and how it is used:

- [https://www.youtube.com/watch?v=VAQMQDJYEZc](https://www.youtube.com/watch?v=VAQMQDJYEZc)
- [https://www.youtube.com/watch?v=FL8awrXz_z8](https://www.youtube.com/watch?v=FL8awrXz_z8)

Please provide your writeup in `.pdf` format, along with the source code and other relevant files packaged in a single zip archive named like `NXSW_firstname_lastname.zip`. If the package is larger than 10Mb please provide a link to download it.

If you have any feedback on this assignment, we're listening.
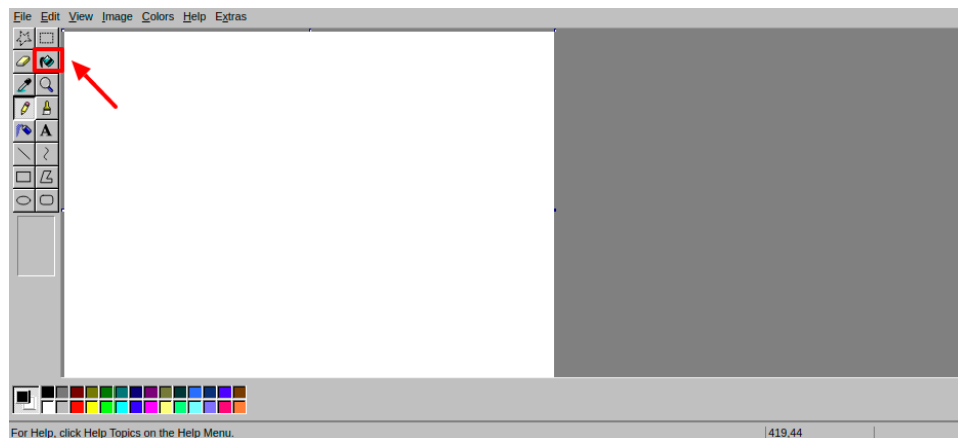
Don't forget, you are not expected to know everything.

Good luck!

# Assignment 1

Your task is to write a function that can be used to implement the bucket tool in MS PAINT:
https://jspaint.app/



Write a function that takes as a parameter:

- 2D array that represent the pixels color of the image an image:
  - assume the colors is an integer from 0 -> 255 (for simplicity)
- A point with coordinates (x, y)
- New color

The function should change the color of the shape that surrounds the point.

In the example below, the square should change color from red to the newColor.

**Example function:**

ChangeColor (Int[][] image, int x, int y, int newColor)
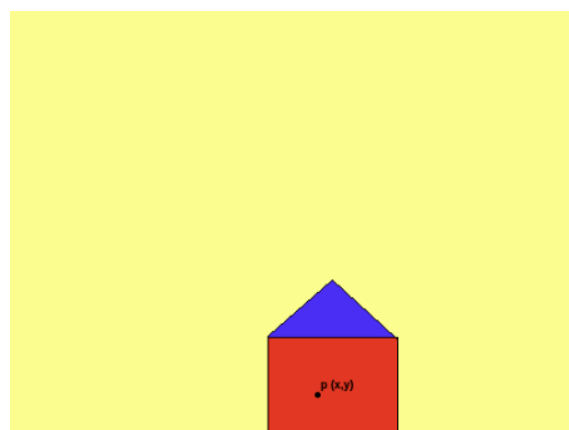
# Assignment 2

You are given a directed graph with N nodes.

Graph can be input via an adjacency matrix or an adjacency list

a) Implement an algorithm that can read an adjacency matrix and builds a graph in memory

b) Implement an algorithm that can read an adjacency list and builds a graph in memory

c) Implement an algorithm that can detect cycles in the graph and can output a warning

d) We consider that the given graph represents dependencies between different nodes.

Implement an algorithm that can output a correct order in which each node needs to be traversed in order to satisfy any dependencies.

```
Input 1 (matrix)              Input 1 (list)

  A B C D E F G               A > B, G

A 0 0 0 0 0 0 1               B > C

B 1 0 0 0 0 0 0               C > D

C 0 1 0 0 0 0 0               D > E, F

D 0 0 1 0 0 0 0               G > A

E 0 0 0 1 0 0 0

F 0 0 0 1 0 0 0               Expected output: Cycle detected

G 1 0 0 0 0 0 0


Input 2 (matrix)              Input 2 (list)

  A B C D E F G               A > B

A 0 0 0 0 0 0 1               B > C

B 1 0 0 0 0 0 0               C > D

C 0 1 0 0 0 0 0               D > E, F

D 0 0 1 0 0 0 0               G > A

E 0 0 0 1 0 0 0

F 0 0 0 1 0 0 0               Expected output: E F D C B A G

G 0 0 0 0 0 0 0
```

# Assignment 3: Device Performance Index

IT Directors, CIOs, and many high-level managers of big companies are some of the main users of our products. These are also the same people that use dashboards and simple KPIs regularly to assess the health of the company's IT infrastructure, as well as to monitor its evolution. In this exercise our goal is to provide them with a global Device Performance Index (DPI), which will be computed for every device. The DPI will be used by our customers to:

- compare the overall device performance across various offices (internal to a customer)
- benchmark overall performance against other customers

For this assignment, you will:

1. Implement a strategy for calculating the DPI, according to the specifications below.
2. Execute the strategy against the sample dataset provided and generate statistics for each client (mean, standard deviation, median, min, max, 25th and 75th percentiles).
3. Tell us what you find from the data in the context of the two use-cases.

The DPI must have values in $[0, 10]$, 0 representing a terrible device and 10 being a top performer.

A sample dataset for computing the DPI is attached in the `device_performance_index.csv` file, where each line contains information about a single device, totaling 9 different customers and 742 offices. Description of the data and units are given in the accompanying `readme` file.

When implementing the code library, please write your code as if it was going into production. You can assume that configuration parameters are passed to your code at initialization time.

## DPI implementation strategy

We first normalize each metric value according to the minimum and maximum values calculated from historical data:

$$X_i^{norm} = \min\left(1, \ \max\left(0, \frac{X_i - X_i^{min}}{X_i^{max} - X_i^{min}}\right)\right)$$

Because all the metrics excepted the `System_Free_Space` indicate worse performance with higher values, we flip them over:

$$X_i^{norm} = 1 - X_i^{norm}, \text{ for } i \in [1,6]$$

Then we simply add them together to obtain the DPI:

$$DPI = \frac{10}{7} \sum_{i=1}^{7} X_i^{norm}$$

Your library needs to include a function to determine the $X_i^{min}$ and $X_i^{max}$ configuration parameters from historical data. Because we want to ignore outliers, $X_i^{min}$ and $X_i^{max}$ will be calculated as the 2nd and 98th percentile of each metric. For the assignment, you will calculate those values from the provided sample dataset, but for production, we will regularly compute it with a dataset containing hundreds of data points for millions of devices.

# Bonus: Alerting Design

In addition to the previous DPI use-cases, many of our clients export the DPI scores and other metrics at regular intervals into their own business intelligence data platform, where they can monitor the evolution of the DPI and act when changes or degradations appear in their dashboards.

We want to expand our product features to cover this monitoring use-case by providing a cross-client cloud solution with:

- integrated DPI time-series data storage, update and retrieval
- automated alert creation through DPI anomaly detection
- real-time notification of users when alerts occur

Your objective is to sketch the high-level design and architecture of a monitoring and alerting system aimed to detect sudden drops of individual DPI scores and changes in around 20 other metrics. The system should leverage cloud services or use horizontally scalable components. Since alerting on a single device can be sensitive to noise, we want to have both alerting at device level as well as alerting at office level, with different sensitivity. The alerts should be accessible through a web front-end, but they should also generate notifications (e-mail, mobile, etc.).

The product and technical requirements for the monitoring and alerting system are the following:

- an alert should be triggered if a significant drop of one of the DPI scores and metrics is detected during the last 10 minutes compared to baselines computed on past data

- alerting should be done at both device and office level

- Solution should scale to 1 billion devices, globally distributed among 1000 clients with an average of 1000 offices/clients (min 10, max 40000)

Along with the design, you should provide us with:

- The assumptions you made

- The existing tools and solutions you would choose and the rationale behind the choice

- The main challenges you foresee

- The tests, or proof-of-concepts, you should build to make sure your solution works in a production environment