

AP+DN7/Analysis Platform+DN7

Q&A

Contents

| | |
|--|---|
| 1 Precautions for use..... | 2 |
| 2 DN7..... | 3 |
| 3 Data Sources..... | 4 |
| 4 Performance..... | 5 |
| 5 Features..... | 6 |
| 6 PCA: Principal Component Analysis..... | 7 |
| 7 Access Management..... | 8 |

| Terminology | Contents |
|-------------------|---|
| AP+DN7 | Abbreviation for AP+DN7. Data analysis infrastructure and tools developed specifically for the use of data in manufacturing processes |
| DN7 | DN7 (Digital Native QC7 Tools) is an analysis tool developed as a QC7 tool for DX. Works on Analysis Platform |
| Edge Server | AP+DN7 app itself or the PC running the app. Currently only recommended for Win 10. It works with Note PCs, etc., but for serious use, it's better to use a PC with better specs. |
| Edge Client | A web terminal on the same network of the Edge Server. Works with tablets such as iPad other than Win 10 |
| Edge | Another name for Edge Client |
| Internal database | A database for buffers created inside PCs running Edge Server. For this reason, choose a PC with plenty of hard disk space for Edge Server. The Edge Client doesn't generate them, so you don't have to worry about capacity. |
| OSS | Open source software (OSS) refers to software that allows users to use, investigate, reuse, modify, extend and redistribute source code for any purpose. |
| OSS License Terms | OSS license terms, which define OSS terms of use based on copyright. Violating this will result in various legal violations and regulations. |
| MIT License | The MIT License is a representative software license originating from the Massachusetts Institute of Technology. It stipulates that software can be handled freely, and that the author or copyright holder has no responsibility whatsoever. |

1 Precautions for Use

| Q | A |
|---|---|
| 1. Can I put it in the One Drive sync folder? | Analytics Platform creates an internal buffer database for fast access, which can cause extreme slowness and network overload if you put it in a folder eligible for One Drive (the web sharing function is designed to reduce network load). Do not use One Drive sync (Avoid places where One Drive is synced, such as My Documents or the desktop, and place it outside of sync.), as this may cause problems. |
| 2. How much free hard disk space do I need? | The internal database can reach several tens of gigabytes when a large amount of process data is read. You do not need to worry about loading CSV etc. once, but please use it on a PC with sufficient free space when loading annual manufacturing data etc. (When the free space on the PC falls below a certain amount, Analysis Platform will automatically stop reading after the warning, but it may be difficult to transfer to another location at that time.). |
| 3. Does this app cost money? Also, can it be used by affiliated companies or overseas bases? | AP+DN7 is open sourced under the MIT License. You can use it for free (your company, subsidiaries, affiliates, suppliers, etc.). There is no limit on the number of people and no limit on the number of introductions. Commercial use is no problem. It can also be used by overseas subsidiaries. It is open source and currently subject to exceptions to the Foreign Exchange and Foreign Trade Act of Japan and the EAR of USA. (However, for greater security for your company, we recommend that you download them directly from GitHub each on your local subsidiaries. That greatly reduces the risk of violating the OSS licensing terms. If you are going to export directly, please do enough OSS license agreement measures and do so at your own risk.) More than 50 languages can be selected for display. |
| 4. Can I modify it and use it? And will it continue to be available even if you stop publishing it? | AP+DN7 may be modified or used under the terms of the MIT license. There are no additional regulations from DENSO, so you are free to modify and use them in accordance with the OSS License Terms. In the future, even if publication is suspended, the published code may still be modified or tested. (The terms and conditions displayed at the beginning of use are almost disclaimer. In the MIT license, it is exempt, but in order to minimize the risk of litigation, the user's intention is confirmed [the application will not launch unless the consent button is pressed].) Please note that modifications not based on the OSS License Terms involve various risks and your company will be held responsible for such modifications. In addition, please note that programs modified in your company may be subject to various regulations, such as the aforementioned export regulations and copyright laws, and should pay deep attention to distribution outside your company. |

2 DN7

Q

A

1. Is it necessary to apply something for DN7 (Digital Native QC tools) to be used?

It is not necessary to apply in particular, and you can download the application from GitHub and use it. (Please refer a link below.)

[Github AnalysisPlatform](#)

3 Data Sources

| Q | A |
|--|--|
| 1. Can I use the connection to the database (DB)? | Please be careful when you use it in this version. There is a risk to be lock the database at worst. |
| 2. What are the points to keep in mind when we try DB connection with Ver. 4.0? | If you are not familiar to database, please avoid using it. If you are professional, please try using it step by step such as trying with a setting with much lower load. |
| 3. No one knows that this setting will not overload the DB. | For Ver. 4.x, we are developing a function for safer data access for large amount of data. If you are not familiar with database, please wait until Ver. 4.x to use DB. In Ver. 5.0, even if the number of clients increases, there is few influences to the load on the referenced DB (Factory DB). |
| 4. What are the points to keep in mind when trying to visualize CSV with Ver. 4.0? | In the case of CSV, there is low risk of using. You feel free to use it. (If you load a csv with a size such as 1GB, the application may die. In that case, if you provide the csv files to us, it might be something better...) |
| 5. How is the date time handled? (On Browser) | All data is handled in UTC within Analysis Platform, but the interface determines the local time according to the locale of the PC, so you can see it in JST from a PC configured in Tokyo. With your PC setup, all the information you operate from the screen should be the time you are using it. |
| 6. How is the date time handled? (In case CSV) | When you import a CSV, the time without time difference refers to the locale of the PC on which the edge server runs, and it is Japan time in Tokyo. For example, if the PC on which the edge server is running is in UK time, the CSV data will be captured as GMT (= UTC), so it will be delayed by 9 hours in Japan. If the PC on which the edge server is running is Japan time, the CSV time is captured as JST. If you are using a regular CSV file, make sure that the time of the PC of the server is the locale in which the data was obtained. |
| 7. How is the date time handled? (In case Database) | When retrieving data from the database, use the locale of the environment where the database is running, if the data contains time differences, or if it does not contain time difference information. Database on Japan Time OS: Treat as Japan Time Database on UK Time OS: Treat as UK time and so on. For example, if you let the edge server in Japan read the value of the database in HUNGARY, the data would be in HUNGARY time. |

4 Performance

| Q | A |
|--|--|
| 1. Do you support real-time display? | Basically, this platform collects data from other database and csv. So the response depends on those systems. If the data source supports real-time update, platform can handle data with a delay of less than 1 minute. (Currently, in this version, in the case of auto update monitoring, selection is limited to 3 minutes cycle. When you submit the button on GUI, you can refer the latest data in the database at that time.) |
| 2. Is it possible to link (JOIN) different databases? | Yes, but you need a common ID for each database. Therefore, if there is a QR code (or other identification) in the product and each system has that information about it, it is possible to link the data. (There are functions such as partial matching and so on.) |
| 3. Will all the data be reloaded from the database for each request? | Basically, the extracted data is reconstructed and accumulated in the internal database to optimize performance. When connecting to the factory database for the first time, the past data is collected at once with bandwidth-limited sequence, but after that, only the differential data is read. Therefore, it is possible to display data at high speed with much lower load on the real factory database. |
| 4. Is it possible to link with other systems? | It can also be linked to ERP, MES, logistics DB, and etc. (common ID is required). When linking with other systems, due to issues such as database load, it is necessary to give due consideration to preparations and collaboration tests with the other system administrator or maintainer vendors. |
| 5. It feels quite dangerous to JOIN to an unprotected (?) database. | That's right. This platform already has some algorithm and architecture to reduce load of database, but still, it is a risk. Therefore, please do not teach the database access information to those who have no knowledge of database for the time being. |

5 Function

| Q | A |
|--|---|
| 1. I wonder why the product number is an integer ... | If only numbers that do not contain the character string like [A-Za-z-! #] are included, they are recognized as integers. In that case, please correct it manually. |
| 2. Is it possible to visualize category values? | Yes, you can. The category functions (StP: stratified plot, PCP: parallel category plot) are also available. |
| 3. Can we remove outliers in time series data, automatically? | Currently it is only by manual. We plan to prepare various type of filters and outlier estimation algorithms in the future. |
| 4. Some CSV files are not read even I add some. | In this version, additional reading is performed only when the latest file name is detected. For example, if the file name has a date, it will not be read even if an old date file is added later. In that case, all the files will be read by resetting the data and reading them again. |
| 5. Are additional data in CSV available? | Once you have checked each plot for automatic updates, you can place a new CSV in the data source folder and it will be read sequentially to update the plot. Therefore, it can be used as an analysis ANDON in the process. In that case, you can change the file name to a permutation number, such as the date and time, so that it reads smoothly. In cases where the equipment is logging in CSV, for example, the generated CSV can be sequentially transferred to the data source folder so that new data can always be displayed. Note, however, that CSVs with different column structures can not be read. |
| 6. What kind of logic is using for outlier recognition? Only when the threshold is exceeded? | In the current version, PCA can use outlier removal (the removal using robust statistical methods, not threshold judgment) to teacher data (it is set as default due to the characteristics of PCA). We plan to add more options in the future. |
| 7. Can we use functions such as threshold judgment and/or QQ plot? | Currently it is not implemented. We will consider introducing those functions in the future depending on the needs. |
| 8. Do I have to install this software on every PC in the factory? | The stand-alone client currently used in Ver. 4.0 is called Edge Server and has a Web service function, so it is possible to refer to data from other PCs / tablets that are external terminals (Edge Client) in the same network. If you can use Chrome on Edge, you do not need to install this app. In addition, Edge does not have to be a Windows terminal. You can use it from Edge by changing the "localhost" of the address to the IP address of Edge Server. |

6 PCA: Principal Component Analysis

| Q | A |
|---|--|
| 1. How do we use PCA? | The comparison source data (teacher data) is regarded as a normal state. If it deviates from the normal state, the index such as T2 or Q (center plots) increases. You can find product ID processed under different (unusual) conditions by watching the reaction of T2 and Q. (for example, if the top item in the plot of 'T2 contribution' {top right plot} is the stage temperature with T2 reaction {top center plot}, we can guess that the product processed at the different stage temperature from the usual.) |
| 2. How do we use the contribution graph in PCA? | Indicators such as T2 and Q (center plots) react by deviating from the normal state. In the contribution plots, items that are most likely to be related to the deviation are displayed in order. So, when you find reaction of T2 and/or Q, large value items in the contribution plot (plots on the right) are considered suspicious. |

7 Access management

| Q | A |
|--|--|
| 1. Do you have settings such as user permissions? | We are planning it on the development roadmap, and are considering specifications, but at this moment, we are developing the main functions first. Currently, data can be referenced from all terminals that can access Edge Server (standalone software of Ver. 4.0), so please be careful about information management such as IP address of Edge Server. |
| 2. When will the access control function be available? | The access management function is planned to be supported in Ver. 5.0 or later where the centralized management function of Edge Server is implemented. |