Zusammenfassung:

Das Executeable Importierung dient zur Bereinigung von erhaltenen Messdaten im CSV-Format. Die wichtigsten Schritte hierbei sind das Entfernen von sogenannten Info-Zeile, diese durch eine Raute markierten Zeilen befinden sich nur zum Debugging im File und sind für die Zwecke dieser Arbeit irrelevant, somit werden sie entfernt, gleich verhält es sich mit den Zahlen am Beginn jeder validen Zeile und das Erkennen Fehlerhafter Zeilen in jeglicher Hinsicht. Die bereinigten Daten werden in ein File gleichen Namens, aber mit dem Zusatz \_result und optional eine Influx-Datenbank geschrieben. Die Info-Zeilen, Fehlerhaften Zeilen und zu großen Zeitabstände zwischen zwei Zeilen werden in einem File mit dem Zusatz \_info vermerkt.

Arten von Zeilen

Im Kontext dieser Arbeit werden Zeilen eines Files in drei Typen eingeteilt:

Info-Zeilen: Beginnen mit einem # und bestehen aus einer einzigen Spalte, sie beinhalten Statusmeldungen des Messgeräts, wie zum Beispiel die Qualität der GSM-Verbindung, ob ein Neustart durchgeführt wurde, et cetera.

Valide Zeilen: Diese entsprechen dem ermittelten Format, dieses wird anerkannt, wenn eine zwei Drittel Mehrheit der ersten 20 Zeilen, ausgenommen sind Info-Zeilen, formatmäßig übereinstimmen. Sie bestehen aus einer Zahl im Bereich 0-60, der Geräteidentifikationsnummer, dem Monat, dem Tag, der Stunde, der Minute, der Sekunde, einem bis acht Messkanälen und einem Diagnosewert.

Fehlerhafte Zeilen: Jegliche Zeilen, welche weder den Vorgaben der Info-Zeile, noch der validen Zeile entsprechen. Durch die umfangreichen Überprüfungen ist sichergestellt, dass neben Zeilen mit Werten auch welche mit dem Wert None als Fehler erkannt werden und keine Exception zur Folge haben.

Programmaufbau:

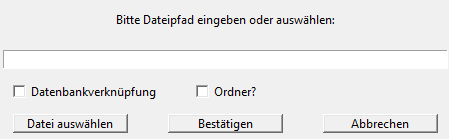
Global werden Arrays für die Info-Zeilen, die Fehlerhaften Zeilen, zu große Zeitabstände und für die Übertragung in die Datenbank deklariert, dies benötigt für die Ermittlung des Wochentags noch eine Formatvorlage.

Die Verarbeitung läuft folgendermaßen ab:

Das File wird geöffnet und eine \_result File erzeugt, das Format wird wie obenstehend erklärt ermittelt, ebenso die Zeitabstände zwischen den Zeilen. Für jede Zeile des Files wird das Format, sofern es vorher bestimmt werden konnte überprüft, stimmt diese nicht überein wird die Zeit im Error Array vermerkt. Im nächsten Schritt wird nach der gleichen Systematik der Zeitabstand zwischen der letzten validen Zeile und der aktuellen Zeile überprüft, Fehler dieser Art werden in einem anderen Array gespeichert, ebenso die Info-Zeilen. Sind alle Überprüfungen abgeschlossen wird die Zeile ohne die Zahl der ersten Zeile in das \_result File geschrieben, bei Datenbankanbindung wird die Zeile mit formatiertem Datum und Angabe des Wochentages in Form einer Zahl von null bis sechs dem betreffenden Array angehängt und die aktuelle Zeile als vorherige Zeile gesetzt. Nach Ende der Schleife werden gegebenenfalls die Daten der Datenbank übermittelt und das Array geleert, was bei Verarbeitung ganzer Ordner essenziell ist, da es Ansonsten unweigerlich zu einem Fehler der Datenbank aufgrund zu großer Datensätze kommt. Abschließend wird noch das Info File erstellt und befüllt und die Verbindung zur Datenbank geschlossen.

Präsentation dem Anwender Gegenüber:

Startfenster: Auswählen ob eine Influx Datenbank vorhanden ist oder nicht. Will man den gesamten Ordner bereinigen oder nur ein File. Auf Datei auswählen klicken, File bzw. Ordner auswählen und bestätigen.



Im Hintergrund wird das File bereinigt und gegebenenfalls in die Datenbank übertragen.

Nach Abschluss erscheint ein Fenster mit einer Entsprechenden Nachricht und im Verzeichnis sind die beiden neuen Files vorhanden.