

An aerial view of an airport tarmac. A large white airplane is parked at a gate. Several white delivery vans with the 'DataFly' logo are parked on the tarmac. Ground crew members in high-visibility vests are visible. The airport terminal building is in the background.

Zwischenstand

Projektrealisierung - DataFly

Aref Hasan, Nik Yakovlev, Christian Schmid, Niklas Scholz, Luca Mohr

Agenda

Arbeitspakete

GANTT-Chart

Process Mining

Data Engineering

Nächste Schritte

Arbeitspakete

Arbeitspakete

Arbeitspaketbeschreibung	
PSP-Code:	2.16
AP-Bezeichnung:	Datenbereinigung
AP-Verantwortung:	Dev-Team
Voraussetzungen:	<ul style="list-style-type: none">• Vollständige und validierte Daten aus dem vorherigen Arbeitspaket (Datenqualitätsprüfung)• Zugriff auf Bereinigungstools
Ziele:	<ul style="list-style-type: none">• Sicherstellen, dass die Daten korrekt sind• konsistent sind• und frei von Duplikaten sind
Inhalte:	<ul style="list-style-type: none">• Erkennen und Entfernen von Duplikaten• Korrigieren von fehlerhaften
Ergebnisse:	<ul style="list-style-type: none">• Bereinigte Daten ohne Duplikate• Korrekte und konsistente Datensätze
Ressourcen:	<ul style="list-style-type: none">• <u>Datenbereinigungstools</u>• Datenanalysten• Speicherplatz für bereinigte Daten
Geschätzter Aufwand:	50 Arbeitsstunden
Starttermin:	30.05.2024
Endtermin:	31.05.2024

Arbeitspakete

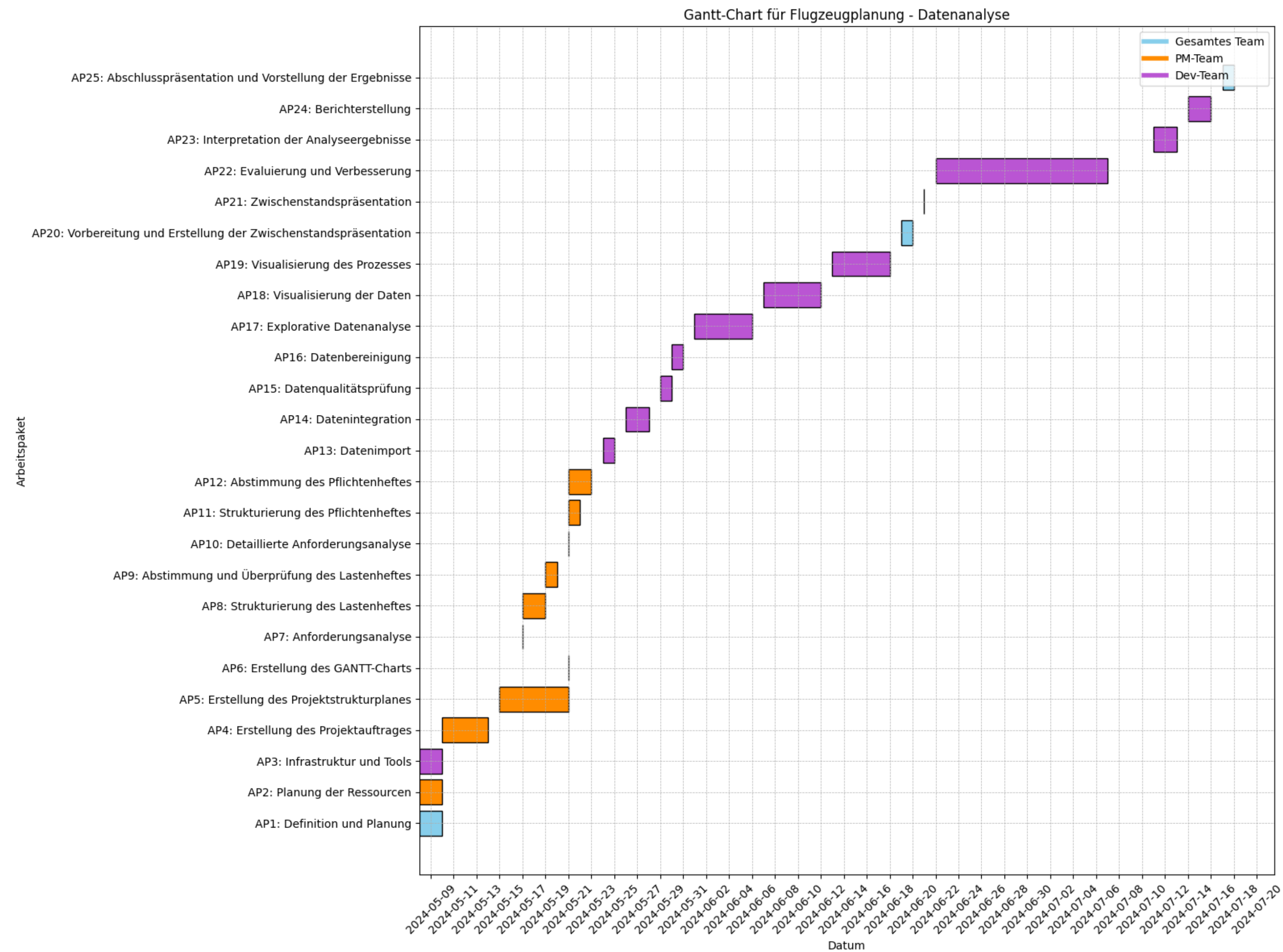
Arbeitspaketbeschreibung	
PSP-Code:	2.19
AP-Bezeichnung:	Visualisierung des Prozesses
AP-Verantwortung:	Dev-Team
Voraussetzungen:	<ul style="list-style-type: none">• Vollständige und validierte Daten aus vorherigen Arbeitspaketen• Zugriff auf Visualisierungstools
Ziele:	<ul style="list-style-type: none">• Entwicklung klarer und verständlicher Visualisierungen der Prozessschritte zur Optimierung der Flugzeugbeladung
Inhalte:	<ul style="list-style-type: none">• Identifikation der Prozessschritte• Entwicklung von Visualisierungskonzepten
Ergebnisse:	<ul style="list-style-type: none">• Klar strukturierte Prozessdiagramme• Interaktive und dynamische Visualisierungen
Ressourcen:	<ul style="list-style-type: none">• Visualisierungstools<ul style="list-style-type: none">◦ Microsoft Visio◦ Tableau• Datenanalysten• Umsetzung der Grafiken
Geschätzter Aufwand:	80 Stunden
Starttermin:	13.06.2024
Endtermin:	18.06.2024

Arbeitspakete

Arbeitspaketbeschreibung	
PSP-Code:	2.6
AP-Bezeichnung:	Erstellung des Gantt-Charts
AP-Verantwortung:	PM-Team
Voraussetzungen:	<ul style="list-style-type: none">• Abgeschlossener Projektstrukturplan• Verfügbarkeit der relevanten Zeitpläne und Meilensteine• Zugang zu den notwendigen Tools zur Erstellung des Gantt-Charts
Ziele:	<ul style="list-style-type: none">• Visualisierung des zeitlichen Ablaufs des Projekts, um die Übersicht über Start- und Endzeitpunkte sowie die Dauer der einzelnen Aktivitäten zu gewährleisten• Unterstützung bei der Planung und Überwachung des Projektfortschritts
Inhalte:	<ul style="list-style-type: none">• Zeitlicher Ablauf des Projekts, inklusive Start- und Endzeitpunkten: Detaillierte Darstellung aller Projektphasen und Arbeitspakete mit ihren jeweiligen Start- und Endterminen.
Ergebnisse:	Ein umfassendes Gantt-Chart, das den gesamten Projektzeitplan abbildet und als zentrales Planungs- und Überwachungsinstrument dient.
Ressourcen:	Projektmanagement-Software / Visualisierungstool
Geschätzter Aufwand:	6 Arbeitsstunden
Starttermin:	21.05.2024
Endtermin:	21.05.2024

GANTT-Chart

GANTT-Chart



Process Mining

Process Mining

- Wesentliche Spalten identifizieren
 - Case ID (`id`)
 - Timestamp (`creation_time`)
 - Activity (`action_name`)
- Ereignisprotokoll (Event Log) erstellen
- Prozessmodell aus dem Event Log erstellen
- Prozessmodell mit geeigneter Visualisierung darstellen

Process Mining

- Daten vorbereitet mit pm4py
- Findung des besten Tools (verschiedene ausprobiert)
 - Visualisierung mit Celonis
 - Akademische Lizenzen beantragt
- Erstellung einer ausführlichen Übersicht mit Celonis

Data Engineering

Data Engineering

Projektrealisierung

- Behebung der korrupten Files
- Konvertierung von CSV-Dateien zu Parquet-Dateien zur Reduktion der Dateigröße
- Verschiedene Formate innerhalb der Entry Details (XML, Tabellenähnliche Formate)
- Gruppierung der Actions nach Formaten
- Einzelne Formate in Tabellenform in Dataframes überführen
- Fehlende Werte innerhalb der Spalten, die nicht als null-Values kodiert waren → künstliches Erzeugen von null-Values
- → zu aufwändiger Ansatz → Konzentration auf Gewichtswerte, die mit Hilfe regulärer Ausdrücke gefunden werden (siehe Ausblick)

Nächste Schritte

Analyse von Plan- und Realwerten

Planwerte (wichtige Actions)

1. **CreateZFWMessageAction:** Erstellen der Gewichtsschätzung (ZFW, Zero Fuel Weight)
2. **EstimateStorePaxDataAction:** Erstellen und Speichern einer Passagierdaten-Schätzung
3. **UpdateEstimatesAction:** Update der Schätzwerte
4. **UpdateCargoMailEstimatesAction:** Update der Fracht- und Post-Daten (geschätzte Werte)
5. **UpdateFuelDataAction:** Update von Treibstoffdaten (geschätzte Werte)

Analyse von Plan- und Realwerten

Realwerte (wichtige Actions)

1. **CalculateWeightAndTrimAction:** Berechnung der Gewichtsverteilung
2. **CargoFinalAction:** Mitteilung der aktuellen Fracht-Daten (finale Werte)
3. **StoreAircraftDataAction:** Speichern von Flugzeugdaten (inkl. realer ZFW)
4. **StorePaxDataAction:** Speichern von Passagierdaten
5. **RampFinalAction:** Schließen eines Fluges durch den Ramp Agent
6. **SendLoadsheetAction:** Senden des Loadsheets
7. **CreateLoadsheetAction:** Erstellen des Loadsheets
8. **UpdateLoadTableAction:** Update von Ladetabellen

Analyse von Plan- und Realwerten

Nächste Schritte

1. Wichtige Gewichtswerte für den finalen Report ermitteln
2. Regex-Funktionen schreiben, um diese Werte zu extrahieren
3. Mögliche Visualisierungen für die Analyse von Plan- und Realwerten heraussuchen

The End



DATAFLY