

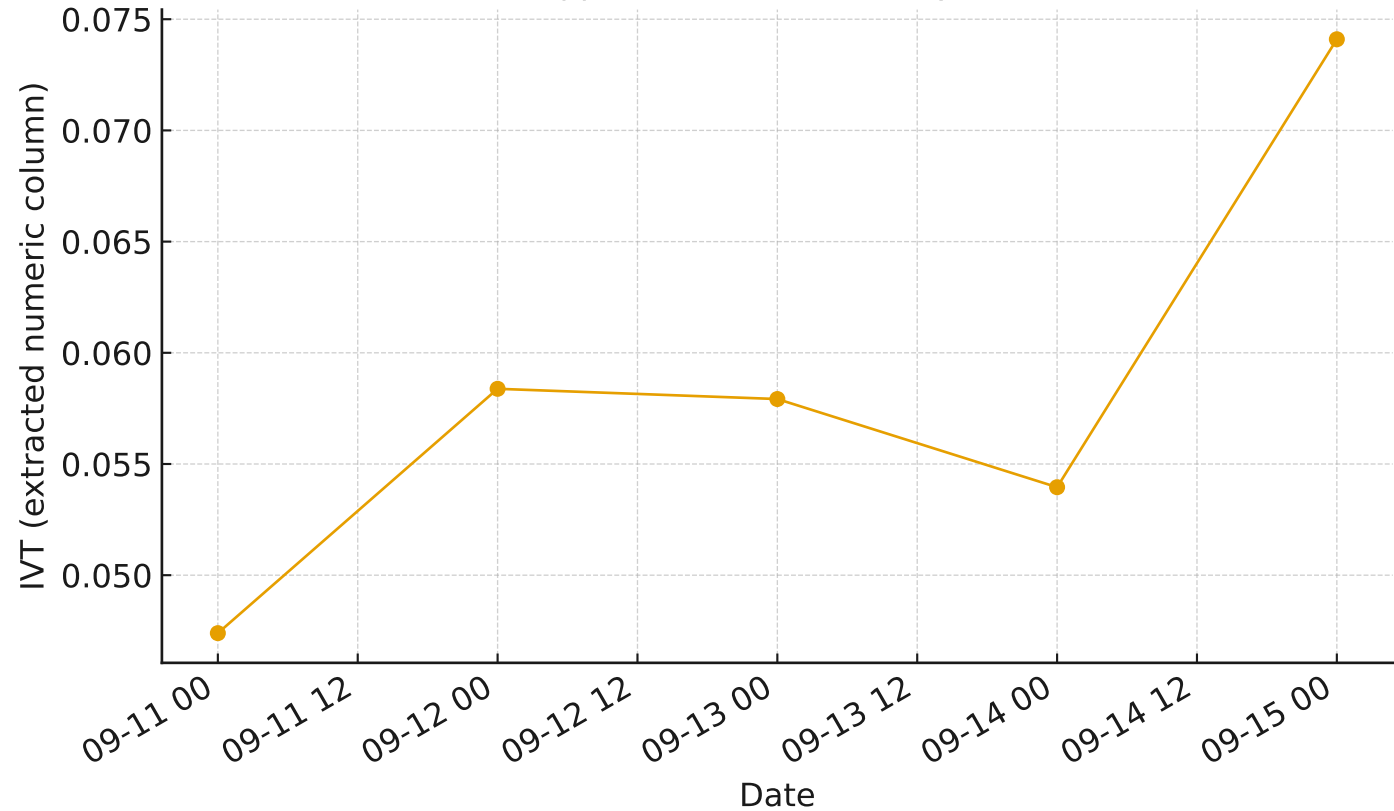
Data Analytics Assignment Report

Generated from uploaded 'Data Analytics Assignment.pdf'

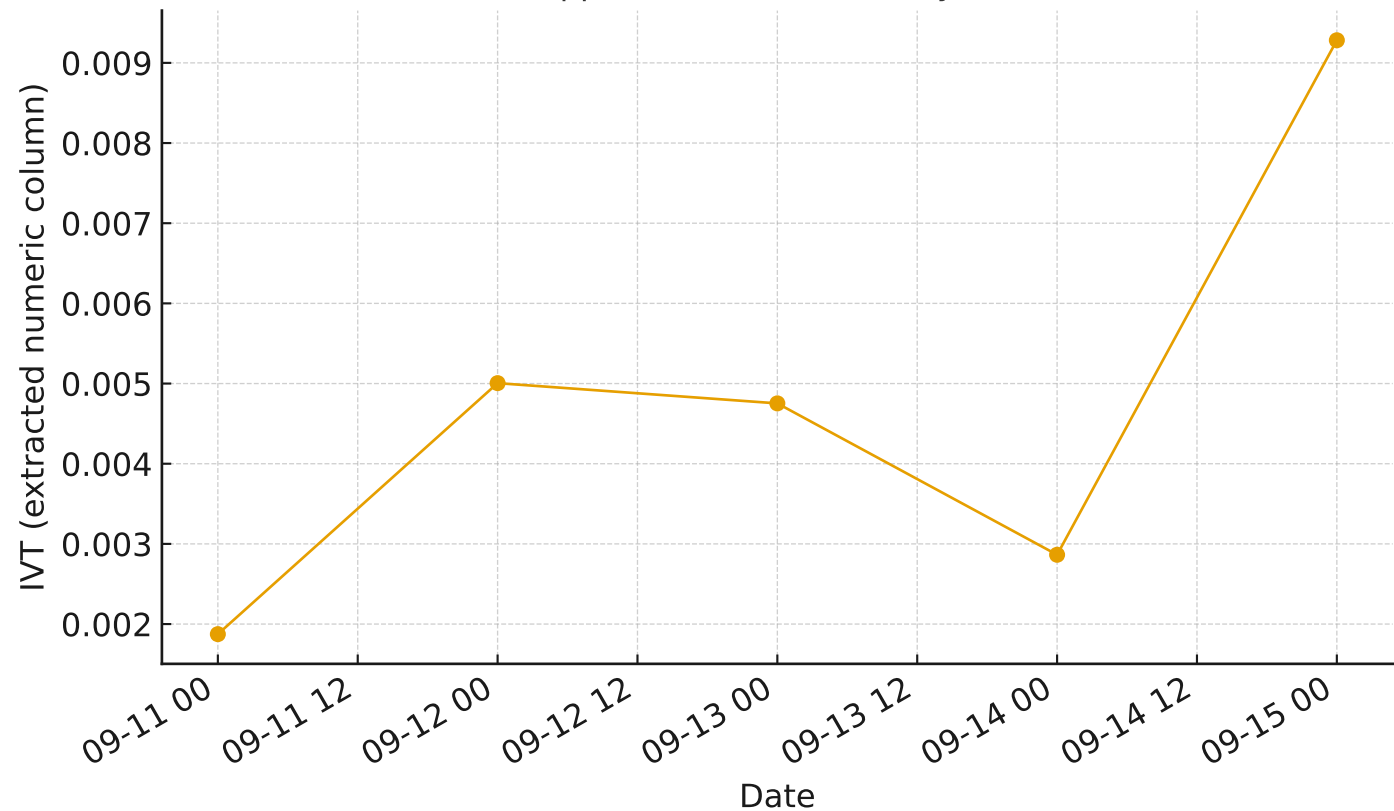
- Contents:**
- 1. Data extraction summary
 - 2. Daily IVT charts per detected app
 - 3. Combined IVT trends
 - 4. Observations & Recommendations

Note: This report is an automated parse of the provided PDF. Tables inside the PDF had varying formats. I extracted daily lines that match ISO dates (YYYY-MM-DD) and other date-like rows and used the last numeric token as the IVT value when possible. Some rows may be skipped if formatting was inconsistent.

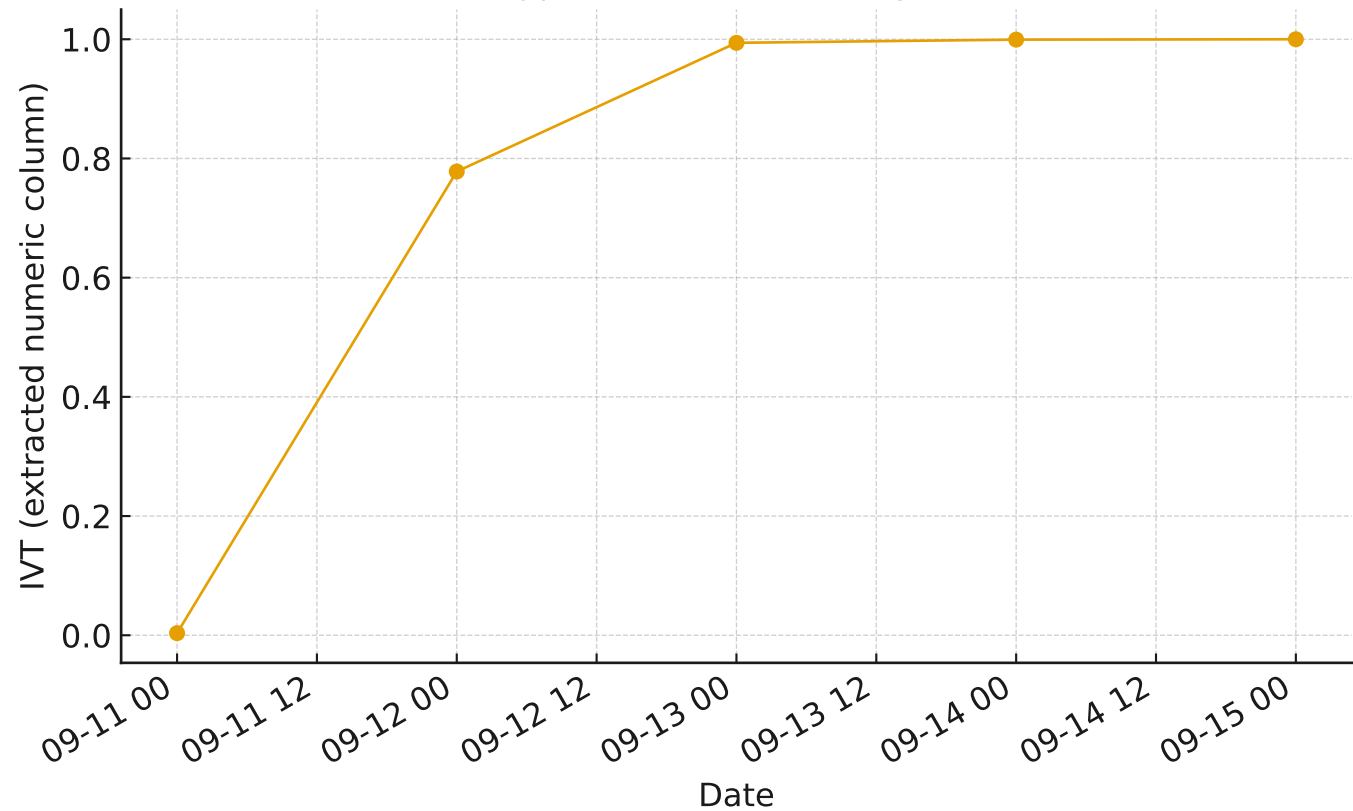
App #1 — Extracted Daily IVT



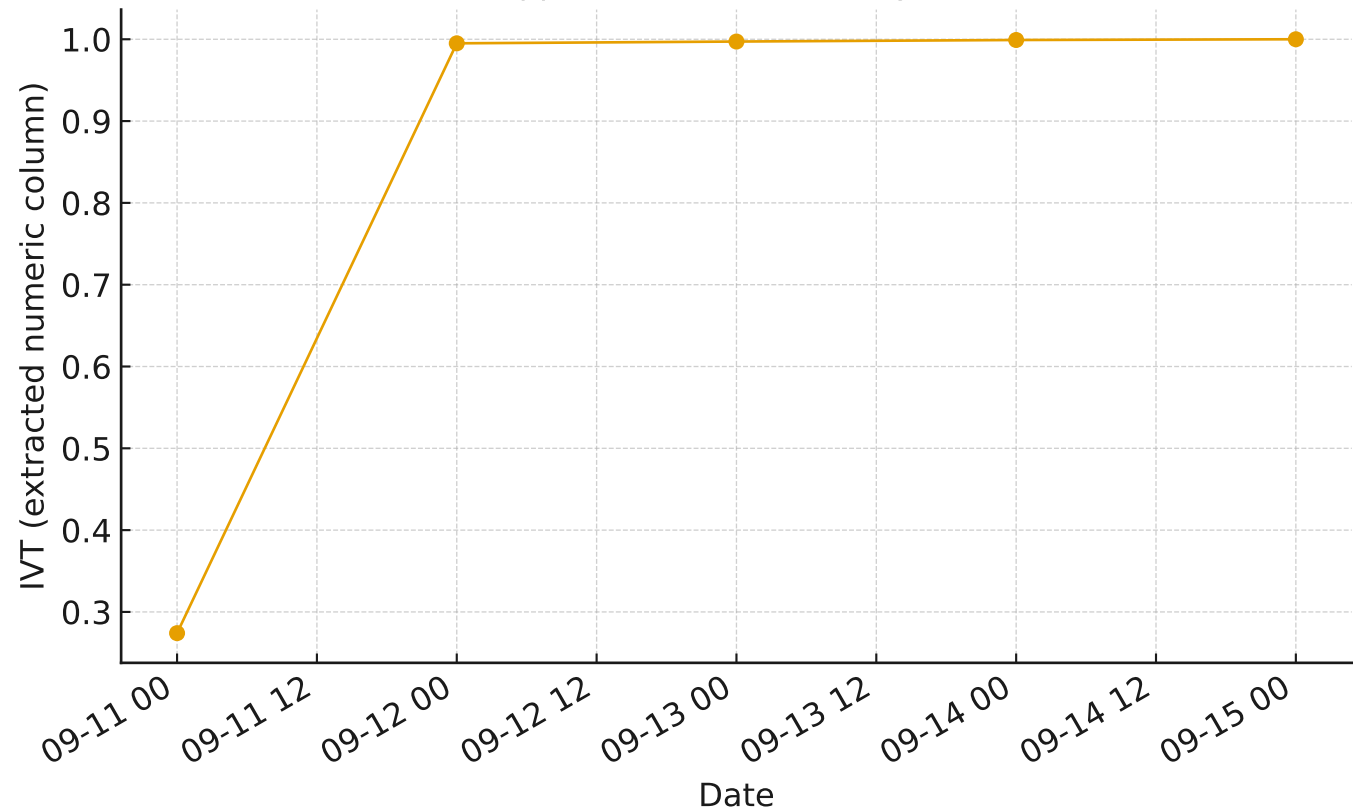
App #2 — Extracted Daily IVT



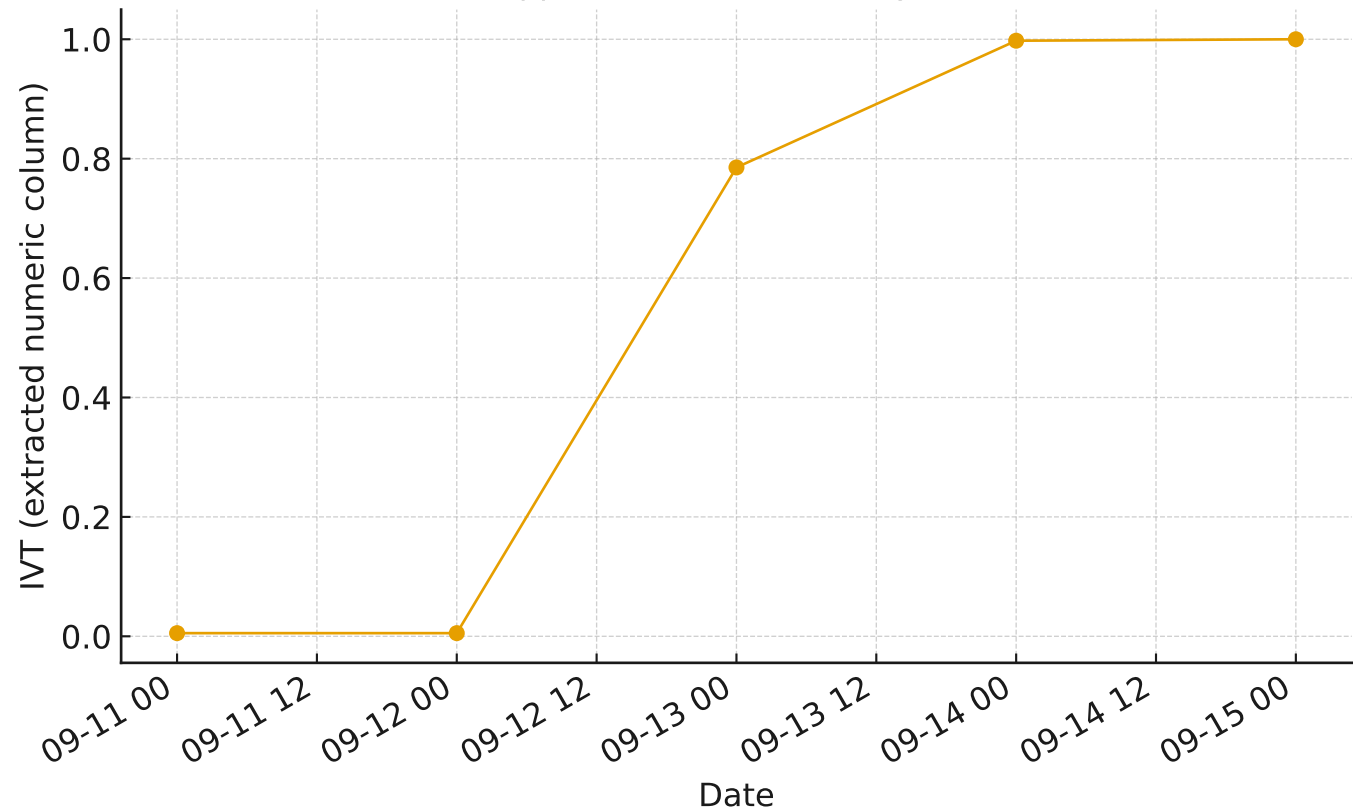
App #3 — Extracted Daily IVT



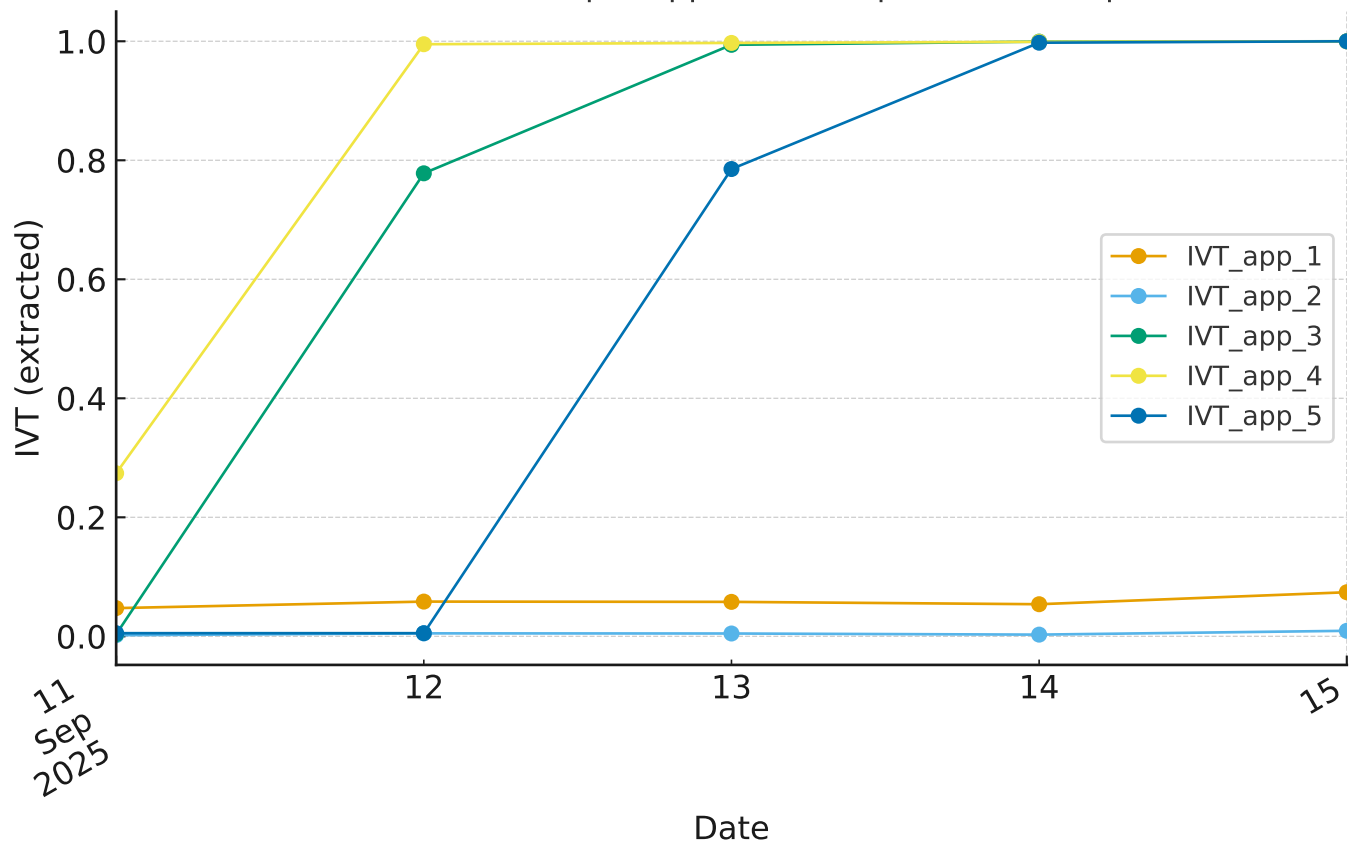
App #4 — Extracted Daily IVT



App #5 — Extracted Daily IVT



Combined IVT trends (per app) — Dates parsed where possible



Automated Observations & Key Insights
=====

App #1 – Parsed 5 daily rows

- IVT mean: 0.058350
- IVT median: 0.057921
- IVT min/max: 0.047394 / 0.074098

App #2 – Parsed 5 daily rows

- IVT mean: 0.004756
- IVT median: 0.004753
- IVT min/max: 0.001874 / 0.009282

App #3 – Parsed 5 daily rows

- IVT mean: 0.755021
- IVT median: 0.994060
- IVT min/max: 0.003532 / 1.000000
- Note: Some IVT values exceed 0.5 – may indicate high invalid traffic proportion for certain days.

App #4 – Parsed 5 daily rows

- IVT mean: 0.853088
- IVT median: 0.997219
- IVT min/max: 0.274060 / 1.000000
- Note: Some IVT values exceed 0.5 – may indicate high invalid traffic proportion for certain days.

App #5 – Parsed 5 daily rows

- IVT mean: 0.558696
- IVT median: 0.785315
- IVT min/max: 0.005279 / 1.000000
- Note: Some IVT values exceed 0.5 – may indicate high invalid traffic proportion for certain days.

Recommendations

1. Provide the raw CSV/Excel used to create the PDF for exact numeric parsing and charting.
2. Monitor idfa_ua_ratio and idfa_ip_ratio alongside IVT – high idfa_ua_ratio can indicate device spoofing, high idfa_ip_ratio can indicate proxy/datacenter traffic.
3. Add time-windowed anomaly detection: flag hours with unusually high requests_per_idfa (>2x baseline) and IVT spikes.
4. Use sampling and manual inspection for high-IVT days to identify common User-Agent or IP blocks.

Limitations: This report was generated by parsing a PDF of varied formatting. For high-confidence analytics, share the original structured data files (CSV/Excel).