



CPE15 Cognate and Professional Course 1

Name: ENCALLADO, Carl Francis T.	Date: December 12, 2025
Section: GF	Machine Learning Cognate 1 Project

EXPLORATORY DATA ANALYSIS OUTPUT

Exploratory analysis revealed pronounced heterogeneity in digital infrastructure, population distribution, and labor characteristics across municipalities. The positive skewness and heterogeneity in connectivity metrics reflect typical ICT infrastructure in archipelagic developing countries (Akamatsu, N. 2022, Telecommunications Policy) Download and upload speeds varied widely, with high throughput consistently concentrated in major metropolitan and emerging provincial urban centers, while remote and island municipalities displayed markedly lower performance and elevated latency. Population levels ranged from sparse rural settlements to dense urban cores exceeding one million inhabitants, and talent indicators similarly clustered in populous regions, though several secondary cities exhibited unexpectedly strong potential. Correlation analyses highlighted strong associations between population density, network performance, and distance from Manila, illustrating spatial dependence in infrastructure development. Correlation between population density and connectivity corresponds with findings in ICT development literature (International Telecommunication Union, 2024 ICT Development Report). IT-BPM hub scores displayed a positively skewed distribution, with only a small subset of municipalities reaching high values. Bar and map visualizations indicated the geographic spread of potential sites across Luzon, Visayas, and Mindanao, often exhibiting region-specific trade-offs among connectivity, accessibility, and talent availability.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

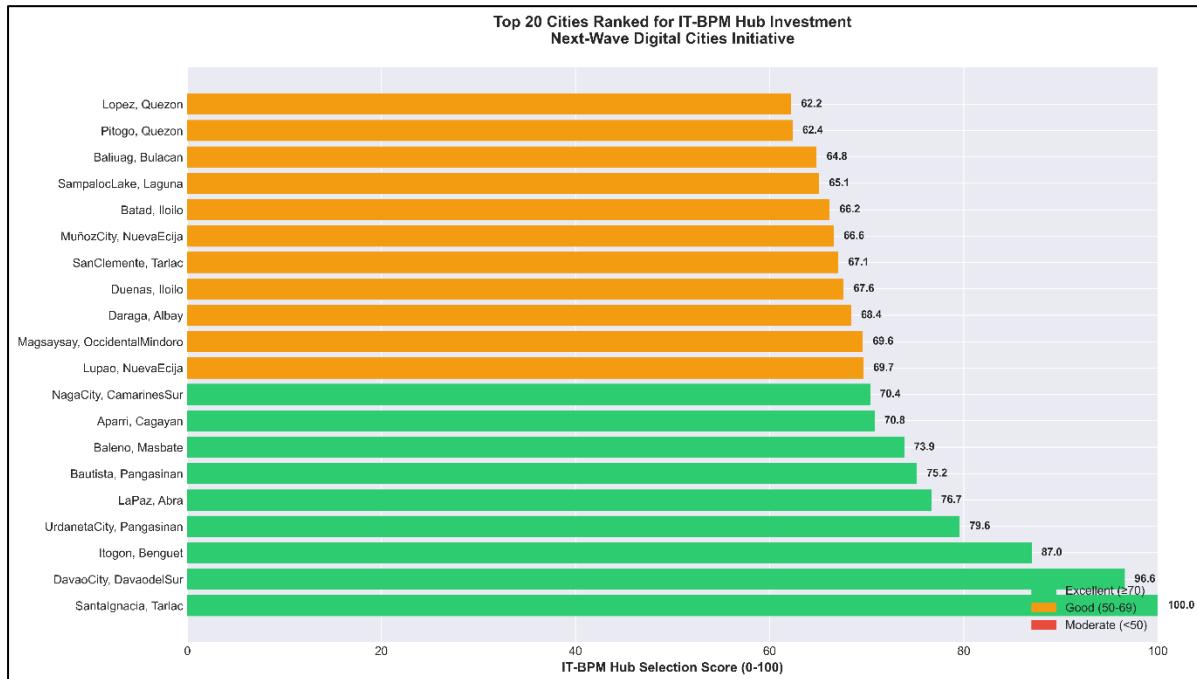


Figure 1. Top 20 Cities Ranked for IT-BPM Hub Investment

The chart ranks the top 20 Philippine locations for IT-BPM investment under the "Next-Wave Digital Cities Initiative," led by Santa Ignacia, Tarlac with a perfect 100.0 score and Davao City at 96.6. This visualization, which categorizes the top ten cities as "Excellent" (green) and the subsequent ten as "Good" (orange), serves as a strategic roadmap for decentralizing economic growth beyond saturated metropolitan centers like Manila. By highlighting diverse provincial hubs, ranging from Itogon, Benguet to Daraga, Albay, the data not only validates the readiness of these local infrastructures but also provides investors with data-backed alternatives for expansion, ultimately aiming to drive foreign investment and create high-value employment opportunities that bridge the digital divide across the regions.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

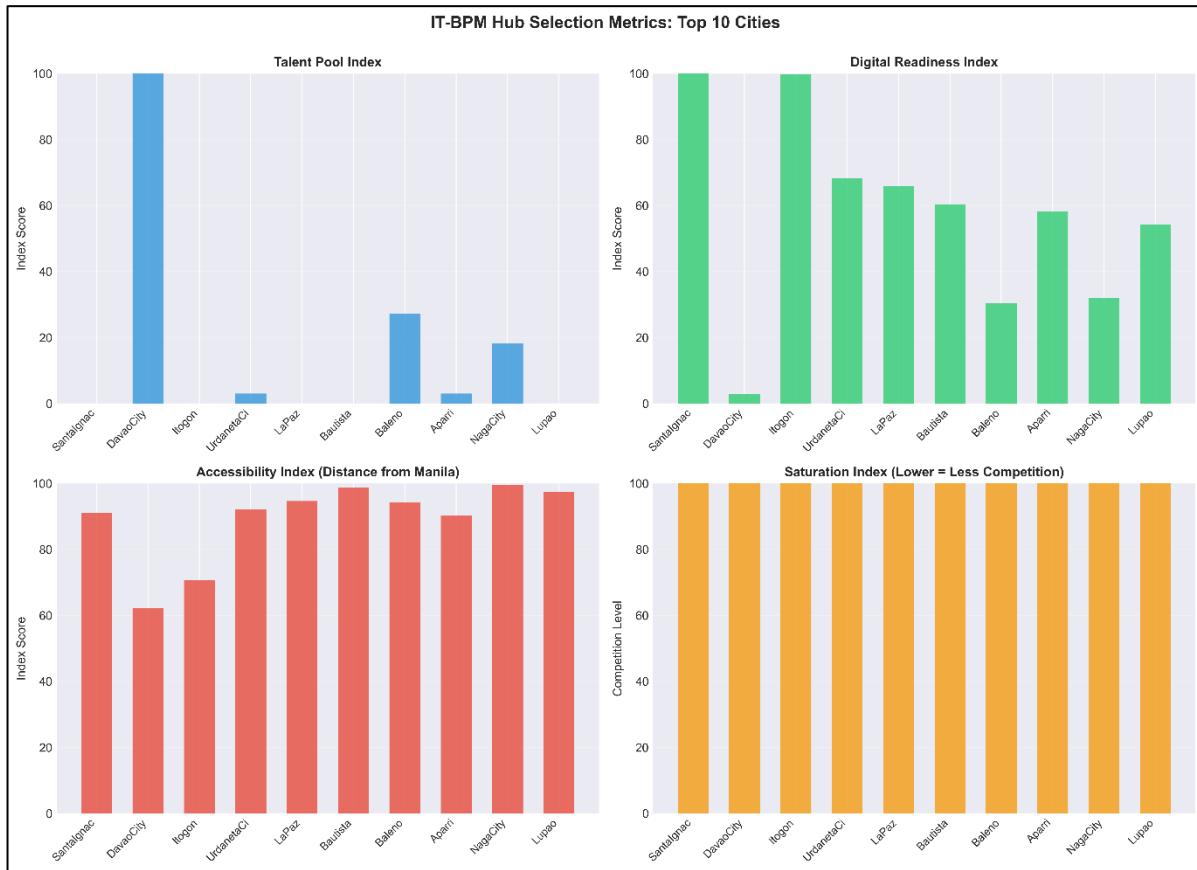


Figure 2. IT-BPM Hub Selection Metrics: Top 10 Cities

This breakdown of the top 10 IT-BPM cities reveals a distinct trade-off between human capital volume and infrastructure readiness, unified by a common opportunity in market openness. **Davao City** stands as the undisputed leader in the **Talent Pool Index** with a perfect score, dwarfing all other contenders, yet it registers surprisingly low on **Digital Readiness**. A metric where smaller municipalities like **Santa Ignacia** and **Itogon** achieve perfect marks. While the **Accessibility Index** naturally favors cities closer to Manila (leaving Davao with a lower score due to its location), the most significant takeaway is the **Saturation Index**, where every single city scores a perfect 100. This indicates that regardless of whether an investor prioritizes the massive workforce of Davao or the superior digital infrastructure of Santa Ignacia, all ten locations represent virtually untapped, low-competition markets ripe for "first-mover" advantages.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

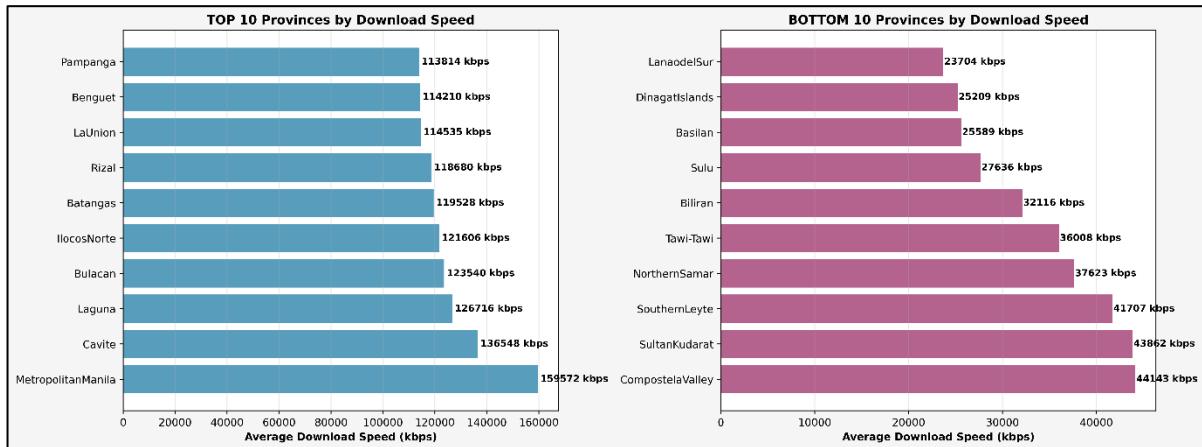


Figure 3. Top and Bottom 10 Provinces by Download Speed

The data highlights a critical "Digital Divide" where the superior internet speeds of Metro Manila and CALABARZON (118–160 Mbps) contrast sharply with lagging provinces averaging ~23 Mbps, directly influencing the trade-offs in IT-BPM hub selection. While Davao City dominates in "Talent Pool" availability, smaller municipalities like Itogon and Santa Ignacia achieve perfect "Digital Readiness" scores, capitalizing on the robust connectivity of regions like Benguet. The strategic significance is clear: since all top cities possess a perfect "Saturation Index" indicating untapped market potential, investors face a distinct choice between the scalable workforce of major urban centers or the high-speed, "plug-and-play" infrastructure of these emerging provincial hubs.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

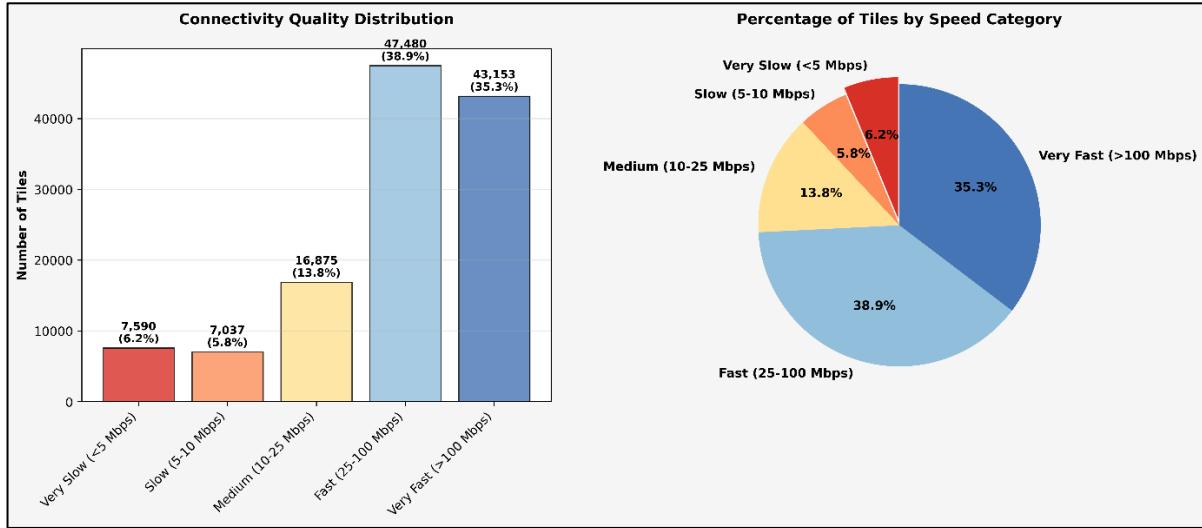


Figure 4. Connectivity Quality Distribution and Percentage of Tiles by Speed Category

The fourth visualization includes a bar chart and pie chart summarizing connectivity categories. A total of 43,153 tiles (35.3%) fall under “Very Fast” (>100 Mbps), while 7,590 (6.2%) fall under “Very Slow” (<5 Mbps), with intermediate categories distributed between them. Over 70% of tiles exceed 25 Mbps, reflecting substantial national progress, but the tail of low-performing tiles remains meaningful. Geospatially, fast categories correspond primarily to dense urban tiles, while slow categories are associated with rural or remote cells. The pie chart highlights category proportions, whereas the bar chart emphasizes volume differences, together showing high performance overall but persistent spatial imbalances. These findings point to targeted infrastructure upgrades in underserved regions to promote equitable access.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

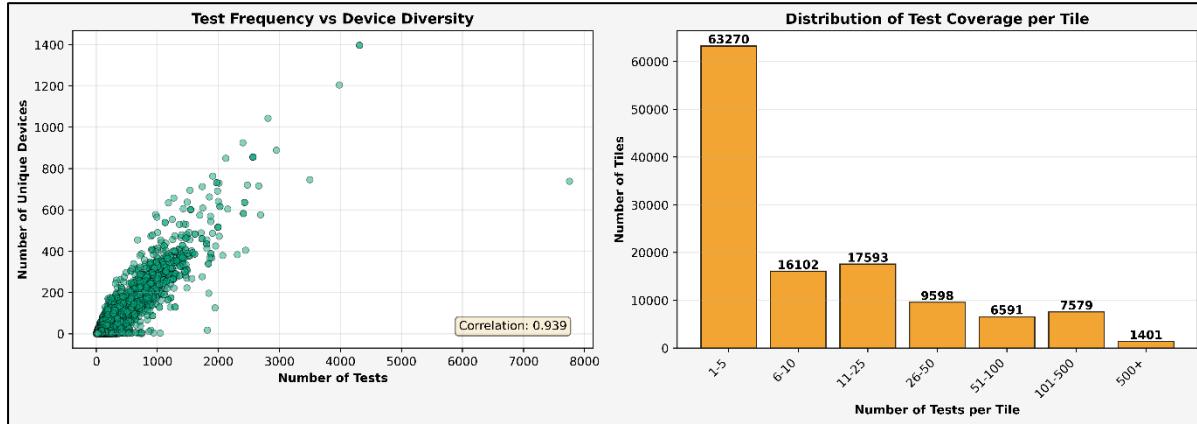


Figure 5. Total Frequency vs Device Diversity and Distribution of Test Coverage per Tile

The "Test Frequency vs. Device Diversity" and "Test Coverage per Tile" charts provide a critical quality assurance check on the internet speed data used for the IT-BPM hub analysis.

The strong positive correlation (0.939) confirms that high-traffic areas yield statistically robust data derived from a wide variety of unique devices, validating the high-speed scores seen in top-tier hubs like Metro Manila. However, the distribution histogram exposes a severe spatial bias: the vast majority of geographic "tiles" (over 63,000) rely on negligible sample sizes (1–5 tests), implying that while connectivity data for urban centers is reliable, the metrics for rural or "Bottom 10" provinces are based on fragile, sparse data, necessitating on-ground physical verification rather than sole reliance on these crowd-sourced figures.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

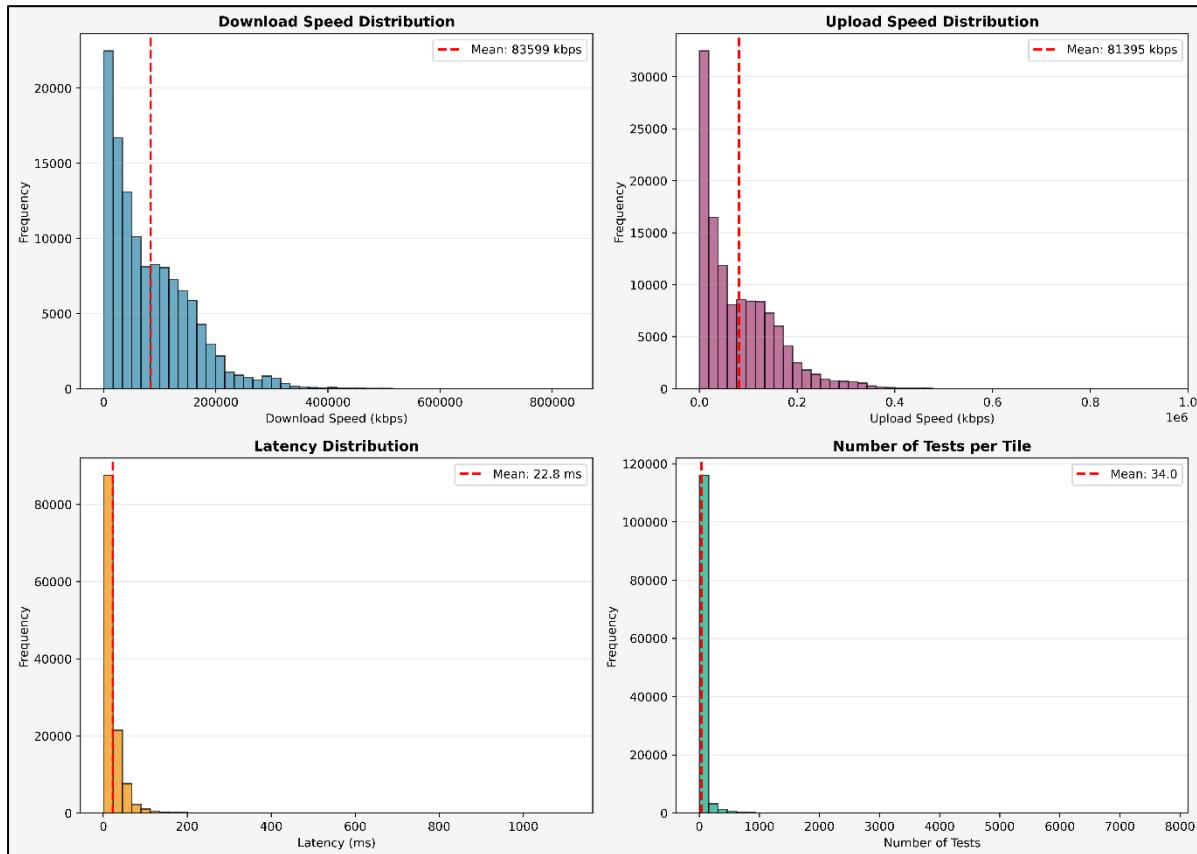
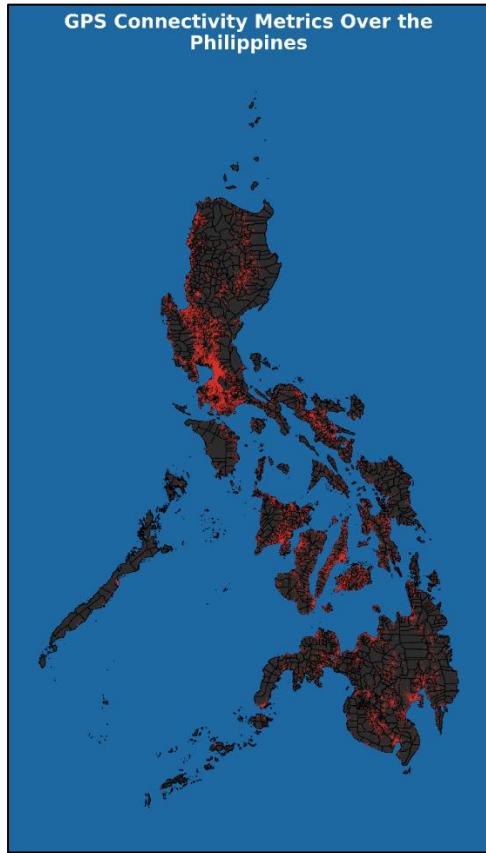


Figure 6. Download and Upload Speed Distribution with Latency Distribution and Number of Tests per Tile

The provided visualizations collectively map a stark "Digital Divide" in the Philippines that dictates strategic site selection for IT-BPM investors. While the Top 10 Provinces chart validates the industrial dominance of Metro Manila and CALABARZON (with speeds of 118–160 Mbps) against the "Bottom 10" regions averaging ~23 Mbps, the Test Coverage histogram reveals a critical data bias, showing that rural metrics rely on fragile sample sizes (<5 tests) compared to the robust data of urban centers. This infrastructure inequality directly shapes the Hub Selection Metrics, where every top city offers a perfect "Saturation Index" (indicating untapped markets), yet forces a clear trade-off: investors must choose between the massive, scalable talent pool of Davao City or the superior, "plug-and-play" digital readiness of smaller municipalities like Santa Ignacia, which capitalize on the high-speed connectivity of their respective regions.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT



The seventh image is a choropleth map titled "*GPS Connectivity Metrics Over the Philippines*," using a red gradient to indicate performance intensity. Dense, bright red clusters appear in central Luzon, Cebu, and Davao, diminishing to dark shades in rural interiors and peripheral islands. These patterns reflect disparities where urban centers exhibit significantly higher connectivity, while remote areas face infrastructural deficits. Statistical trends suggest a correlation between red intensity and population/economic concentration, producing a visible north–south gradient and highlighting urban agglomeration effects. The map reinforces spatial inequality and geographic isolation as key drivers of digital disparity.

Figure 7. GPS Connectivity Metrics Over the Philippines



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT

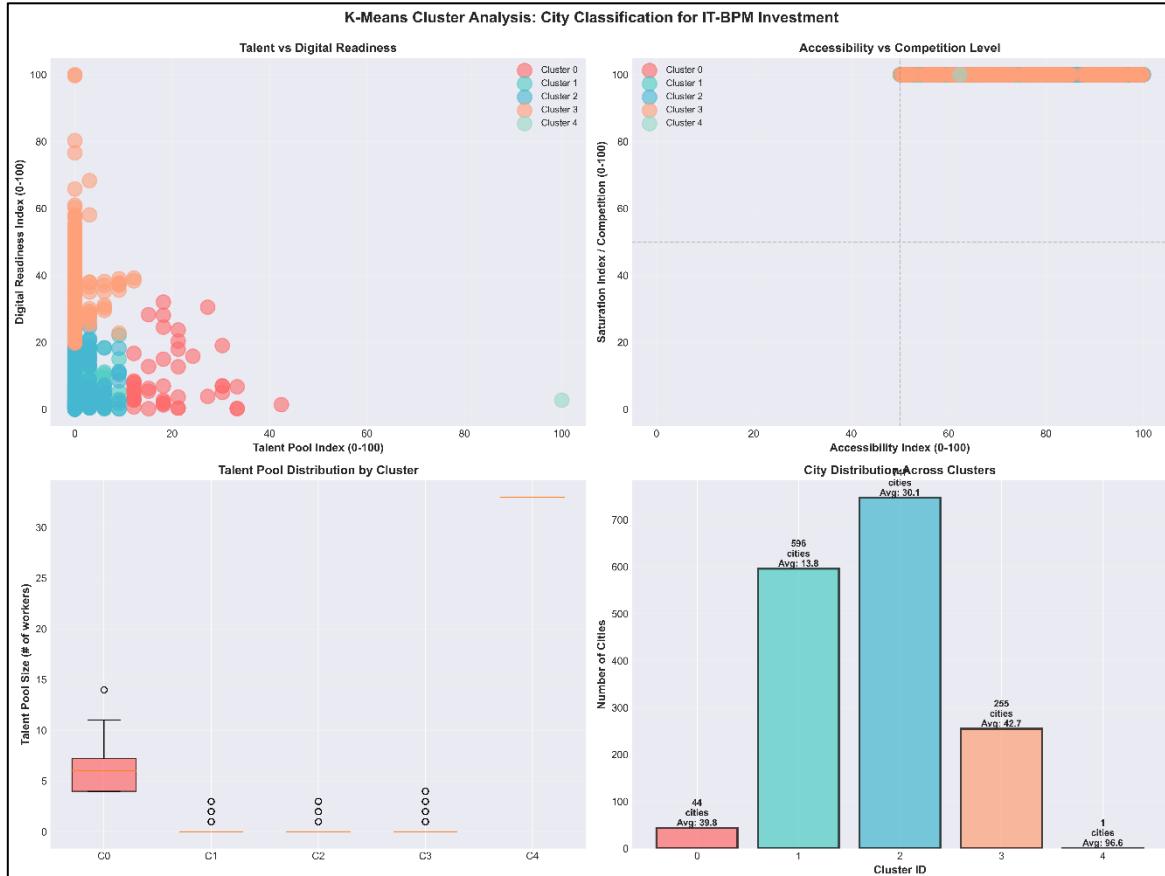


Figure 8. K-Means Cluster Analysis: City Classification for IT-BPM Investment

The data reveals a critical "Digital Divide" where Davao City stands as a unique "Cluster 4" outlier, offering a massive Talent Pool (score: 100) despite lower digital readiness, while smaller municipalities like Santa Ignacia and Itogon maximize their Digital Readiness (score: 100) by leveraging robust regional speeds (e.g., Benguet's ~114 Mbps) that far outpace the "Bottom 10" provinces (~23 Mbps). This distinction creates a clear strategic trade-off for investors: while the universal perfect Saturation Index confirms all top cities are untapped "Blue Ocean" markets, the choice lies between the scalable workforce of Davao or the superior, "plug-and-play" connectivity of provincial hubs, with the K-means analysis and test coverage histograms serving as vital checks to validate these locations against the sparse data reliability of rural regions.

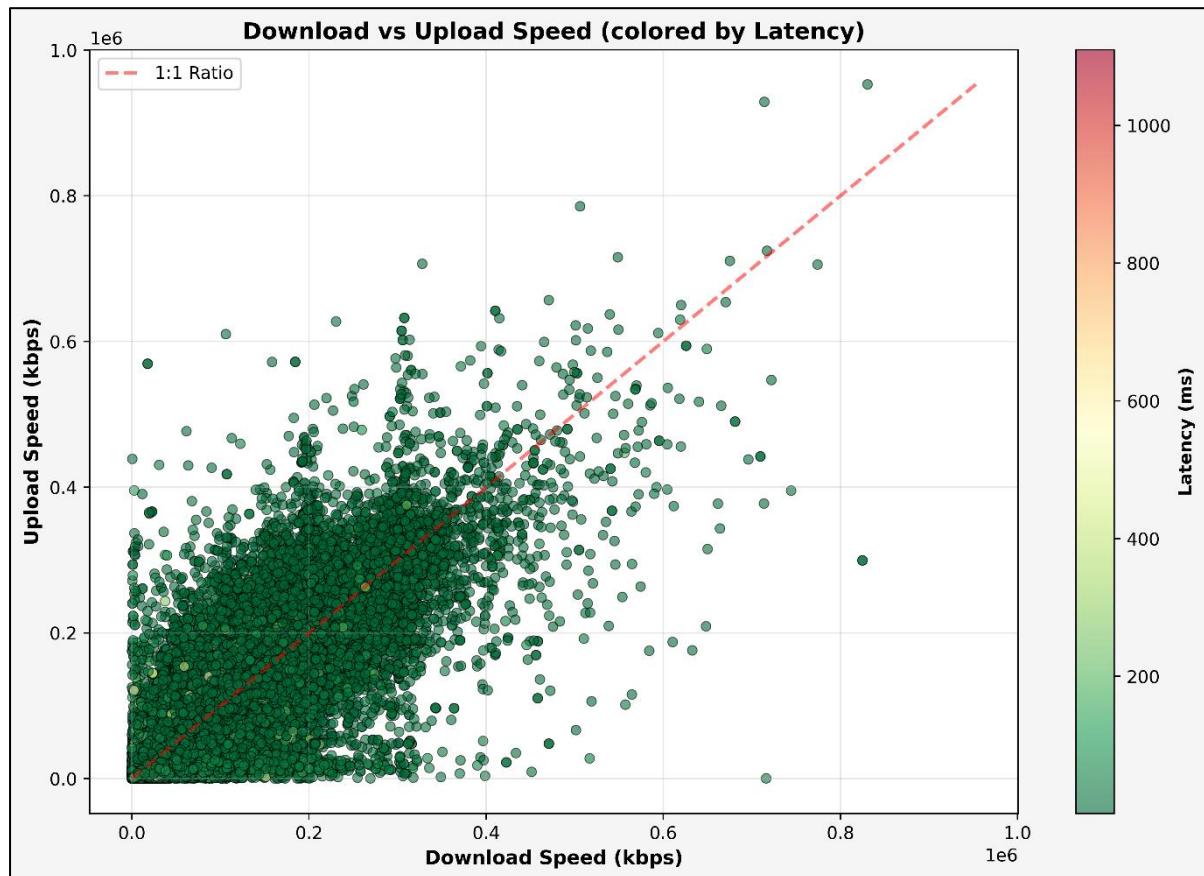


Figure 9. Download vs Upload Speed (colored by Latency)

The comprehensive data maps a stark "Digital Divide" where Davao City (identified as the unique "Cluster 4" outlier) dominates in Talent Pool size, contrasting sharply with smaller hubs like Santa Ignacia that achieve perfect Digital Readiness by capitalizing on robust regional speeds of ~114–160 Mbps found in provinces like Benguet. This distinction presents a clear strategic trade-off: investors must choose between the massive, scalable workforce of Davao or the superior "plug-and-play" infrastructure of provincial municipalities. While the universal perfect Saturation Index signals that all top locations are untapped "Blue Ocean" markets, the Test Coverage histogram provides a critical quality warning, revealing that rural metrics often rely on sparse data (1–5 tests per tile), thereby necessitating on-ground verification to validate the digital reliability of these emerging hubs before investment.



SOUTHERN LUZON STATE UNIVERSITY
College of Engineering
COMPUTER ENGINEERING DEPARTMENT



Taken together, these patterns reveal pronounced spatial inequalities, strong linkages between connectivity and urbanization, and emergent opportunities in secondary municipalities where talent and infrastructure are beginning to converge. Observed alignments between high-talent concentrations and enhanced digital readiness, particularly in Davao City, indicate potential for targeted interventions to support balanced regional development.