Data-centric Modeling of Gainesville Businesses

Michael Elliott, Erik Bredfeldt, Matthew Collins, Renato Figueiredo, Mark Girson,
Amardeep Siglani, Lila Stewart, José A. B. Fortes

Introduction and Goals

In **smart cities**, data collection and analysis enable well informed decisions by city governments and their citizens to create metropolitan areas that are safe, economically sustainable, socially harmonious, and environmentally friendly.

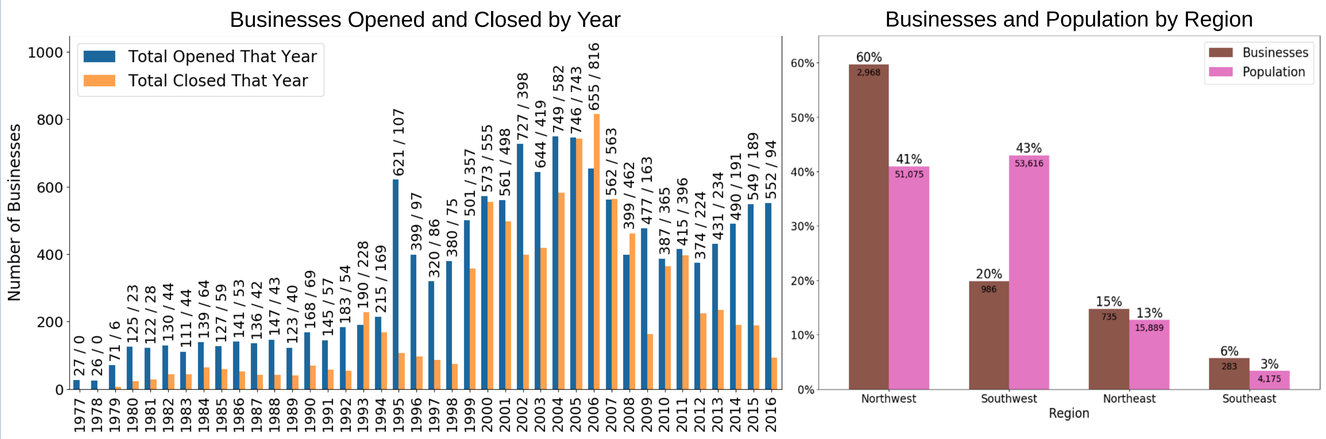
One purpose of this study was to investigate **whether available data can be used** to improve our understanding of how City actions (or lack thereof) impact (positively or negatively) **local business success**.

We also investigate which data should be collected, how to collect and improve that data, and what can be done to facilitate better data access and analysis.

Gainesville City Businesses

As of November 1st, 2017, there were 5528 active businesses in Gainesville.

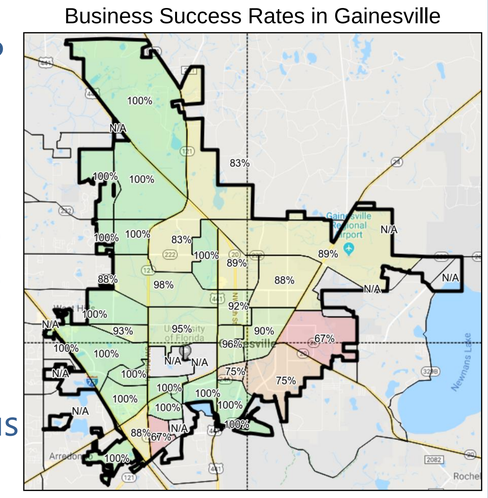
- 22% retail; 19% professional, scientific, and technical services (lawyers, attorneys, tax services, auto repair, laundromats); 18% other services (including barbers, tailors and car cleaning)
- More than half of active businesses are younger than 10 years; oldest businesses are almost 70 years old
- Most businesses (60%) are located in the northwest region of Gainesville

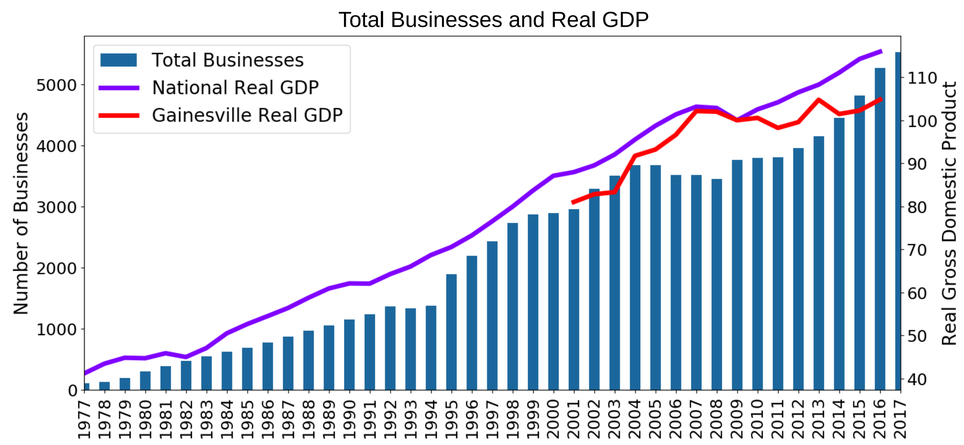


Combined Data and Survey Results Analysis

The survey results (e.g. **business success**) and publicly available data (**business age, electricity consumption, nearby crime, building permits, building code violations,** and **zoning violations,** local and national **Gross Domestic Product** and **Consumer Sentiment Index**) were cross analyzed to identify any existing correlations.

- There is a strong correlation (0.90) between national GDP and the GDP of Gainesville
- There is also a high correlation (about 0.82) between national GDP and business creation in Gainesville
- The 8% of businesses who consider themselves unsuccessful are concentrated in the southeast
- In all but 4 of Gainesville's 44 Census tracts, at least 80% of businesses consider themselves successful



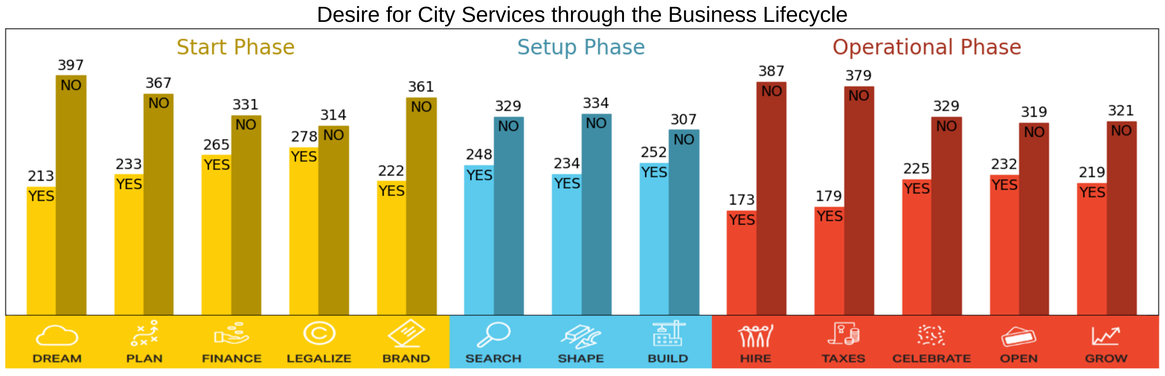


Gainesville City Business Sentiments

Of the 3,821 businesses contacted, 644 responded, giving a 16.9% response rate.

The survey asked business owners whether they desired city services to be available for each of the thirteen identified stages of the business lifecycle*, which are grouped into three phases:

- **Start Phase**: the Department of Doing of the City of Gainesville helps identify resources and expertise to understand the local market and available locations
- **Setup Phase**: the Department of Doing helps business owners understand planning, zoning, and permitting constraints
- **Operational Phase**: the Department of Doing acts as a connector, enabling business owners to identify and source necessary vendors, partners, and services needed to open



Data Quality and Organization

The City of Gainesville provides public access to over 280 online datasets through *data.cityofgainesville.org*, which were supplemented by the survey as well as data from the City, GRU, and federal webpages (Census Bureau, Department of Labor, etc.).

Gainesville's public database should seek to address:

- **Sparse data** – few businesses have data available across all relevant datasets; e.g. 80% of inactive businesses have no type
- **Identifiers** – businesses are currently referenced across datasets only by name, address, or owner, which do not always match
- **Data integrity** – many data entries are faked or erroneous
- **More data** – more data are needed in both quantity and category before meaningful analysis can be done

Conclusions

Analysis of currently available data can be used to glean some useful conclusions, but further research and refined data collection practices are needed to meaningfully model local business success.

There are significant benefits and opportunities to be gained from a City-level data management and analytics framework which:

- **Enables the collection of quality data and cross-dataset queries**
- **Supports a variety of data** that may describe environmental sensing, traffic monitoring, public transportation monitoring, citizen-provided information, personal devices, City-deployed sensors, and operational records from private and public entities

Acknowledgements

Funded by a community research award from the Office of the Senior Vice President and Chief Operating Officer of the University of Florida. Research was done in collaboration with the City of Gainesville Department of Doing and the Bureau for Economic and Business Research at the University of Florida.

* Business lifecycle model inspired by the Blue Ribbon report of the Advisory Committee on Economic Competitiveness to the Gainesville City Commission

PRAGMA 35: October 3-6, 2018
This material is based in part upon work supported by the Office of the Senior Vice-President and Chief Operating Officer of the University of Florida, the National Science Foundation under Grants No. 1550126 and No. 1234983 and the AT&T Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of UF, the National Science Foundation or AT&T Foundation."