On the Feasibility of Data-centric Modeling of Gainesville Businesses

Michael Elliott*, Amardeep Siglani*, Mark Girson*, Erik Bredfeldt[+], Lila Stewart[+], Matthew Collins*, Renato Figueiredo* and José Fortes*
(* University of Florida [+] City of Gainesville)

September 2018

Contents

Sunset over downtown Gainesville, Florida

Foreword

Critical to the success of any city is a clear understanding of the multiple stages of the lifecycle of its businesses, their interactions with the city government, and their success rates. Such understanding enables the city government to attract, foster, and support businesses that serve citizen needs, thus strengthening the city's economy and quality of life. Can we access data from which the desired understanding can be deduced for the City of Gainesville and, if so, how can it be done? The authors of this report, consisting of UF researchers from the Advanced Computing and Information Systems Laboratory and the Bureau of Economic and Business Research and City of Gainesville staff members from the Department of Doing, teamed up to make progress towards addressing these questions.

We are thankful for the funding of this effort provided by a community research award from the Office of the Senior Vice President and Chief Operating Officer of the University of Florida. We also would like to thank the Gainesville business community members who responded to a survey requesting their feedback on their interactions with the City of Gainesville, their assessment of the performance of their businesses and what might have impacted it.

The team responsible for the work reported here consists of UF researchers José Fortes, Renato Figueiredo, Mark Girson, Matthew Collins, Amardeep Siglani and Michael Elliott and City of Gainesville staff members Erik Bredfeldt and Lila Stewart.

José A. B. Fortes
Project Director on behalf of the Project Team
Advanced Computing and Information Systems Lab
Department of Electrical and Computer Engineering, University of Florida

Section I – Introduction

Cities that thrive economically and provide high-quality of life to their citizens depend on information technology to collect and analyze many kinds of data. The ultimate goal is to understand how city services are meeting their purposes, how private businesses are performing, how civic initiatives are being enabled, how city environments are being protected and, in general, how a city is functioning as an integrated whole that serves its citizenry. Increasingly, the objective is to create smart cities where data collection and analysis enables well informed decisions by city governments and citizens to create metropolitan areas that are safe, economically sustainable, socially harmonious and environmentally friendly.

The City of Gainesville (hereafter abbreviated as the City) collects several kinds of data to support its administrative duties, to enforce regulations, protect citizens and manage services. We investigated the possibility of analyzing these and other data to understand the performance and distribution of businesses in Gainesville.

Business lifecycle model

The Department of Doing of the City of Gainesville is creating mechanisms to support businesses at the different stages of development by providing tools and personal interactions that inform, guide, and enhance efficiency of business actions and operations. We adopted a business lifecycle model inspired by the Blue Ribbon report of the Advisory Committee on Economic Competitiveness to the Gainesville City Commission[1]. This model considers the following three phases of a business (in chronologic order), each with several of a total of thirteen stages:

- Start Phase: has five stages -- Dream, Plan, Finance, Legalize and Brand – during which the Department of Doing can help identify resources and expertise to understand the local market, available locations and spaces.
- Setup Phase: includes three stages -- Search, Shape and Build – during which the Department of Doing can help business owners understand relevant planning, zoning, and permitting constraints.
- Operational Phase: includes five recurring stages -- Hire, Taxes, Celebrate, Open and Grow – during which the Department of Doing acts as a connector, enabling businesses to identify vendors, partners, and services.

[1] http://www.cityofgainesville.org/Portals/0/clerk/CityComm/BlueRibbonReport.pdf

Main goals

One of the purposes of this study was to investigate whether available data can be used to improve our understanding of how the City actions (or lack thereof) impact (positively or negatively) the efficiency and success of businesses throughout their lifecycle. Parts of this broad objective include understanding whether City-business interactions are hindering or facilitating business creation and operation and whether the City can gain insights into the reasons for success or failure of businesses (e.g. location) to help new businesses succeed and existing businesses grow and expand.

The approach tested in this project to support the above-mentioned objective is to use data from multiple sources to infer the efficacy of City-business interactions and identify business health indicators. Data were collected from multiple sources, including mainly the ones discussed in Section 2. While it was not the goal of this study to do an in-depth analysis of the economics aspects of Gainesville businesses, the results of this approach may also be of use to those who might be interested in undertaking such analysis in the future.

A related purpose of this project was to investigate what desirable data should be collected, how to collect them, how to improve them, and how to enable data access and analysis. As additional results, this study includes recommendations for improvements in collecting and managing data to complement existing efforts by the City of Gainesville, which already provides open access to over 250 sets of data under the categories of Economic Development & Redevelopment, Environment & Energy, Governance, Human Potential, Infrastructure & Transportation, Public Safety, and Quality of Life.

Methodology

Reflecting the data-centric nature of this study, we started by conducting a discovery process to identify datasets that are available and potentially could inform us of the nature, distribution, environment and performance of businesses in Gainesville. Datasets were obtained from publicly available government sources at the City, State and Federal levels. We mined these datasets for information related to the types of businesses, their geographical distribution, their interactions with the City and factors that could impact business performance. To supplement these sources and get additional information that is not available from public datasets, we conducted a survey of Gainesville businesses to get data on their

performance and sentiment regarding their interactions with the City of Gainesville. We then looked for correlations between performance information reported by survey respondents and data from other datasets. Additional analyses were conducted on the available datasets in order to discover facts, trends and patterns that could be further used to infer business performance and factors impacting it.

All the datasets collected for this study are summarized below. All data and conclusions reported in this study are based on these datasets.

Data availability and sources

Different sources of data are publicly accessible. One source is the Gainesville City Open Data web site[2] -- Table 1 provides a summary of the available datasets that were considered in this study and some of their characteristics.

Table 1 – Datasets from Gainesville Open Data Web Site

Dataset Name	Type of Data	Relevant Fields
Active Businesses[3]	Data about currently active businesses (7,044 records, 1948-2017)	Business Name, Address, Type and Start Date
Building Permits[4]	Data about building permit requests by businesses for repairs or expansion in Gainesville (16,795 records, 2004-2016)	Permit, Parcel No., Permit Type and Permit date
Crime Incidents[5]	Data about crimes in Gainesville (123,000 records, 2011-2017)	Location and Incident Type
GRU Customer Electricity Consumption[6]	Data about the electricity consumption by consumers in Gainesville (7 million records, 2012-2017)	Service Address and Consumption

[2] Gainesville City Open Data: https://data.cityofgainesville.org/
[3] Active Businesses: https://data.cityofgainesville.org/Strong-Economy/Active-Businesses/hk2b-em59
[4] Building Permits: https://data.cityofgainesville.org/Strong-Economy/Building-Permits/p798-x3nx
[5] Crime Incidents: https://data.cityofgainesville.org/Community-Model/Crime-Incidents-2011-Present/gvua-xt9q
[6] GRU Electricity Consumption: https://data.cityofgainesville.org/Better-Future/GRU-Customer-Electric-Consumption/gk3k-9435

The data shown in Table 2 were also provided by the City and Gainesville Regional Utilities (GRU) on request. Additional publicly available data are listed in Table 3. They include anonymized aggregate data related to the Gainesville business survey discussed in Section III.

Table 2 – Datasets requested from GRU and the City of Gainesville

Dataset Name	Type of Data	Relevant Fields
Commercial Building Code Violations[7]	Data on building code violations by businesses in Gainesville (550 records, 2010-2017)	Violation, Infraction Date, Parcel and Address
Commercial Zoning Violations[8]	Data on zoning code violations by businesses in Gainesville (543 records, 2010-2017)	Violation, Infraction date, Parcel and Address
Business Lifespans[9]	Lifespan of both active and inactive businesses in Gainesville (7,088 active business records, 8,651 inactive business records, 1948-2017)	Business Name, Start Date, End Date and Business Type
GRU Utility Consumption[10] (several sets)	Data about utilities consumption of consumers in Gainesville (2006-2017)	Premise, Install, Contract, Contract Account, Business Partner and Consumption
Parcel Data[11]	Parcel data for businesses in Alachua County (148,277 records)	Parcel, Full Address, Latitude and Longitude

[7] Code Violations: Provided by City Code Enforcement on request.
[8] Zoning Violations: Provided by City Code Enforcement on request.
[9] Business Lifespan: Provided by GRU on request.
[10] GRU Utility Consumption: Provided by GRU on request.
[11] Parcel Data: "AddressPoints" dataset at http://data-acpa.opendata.arcgis.com/datasets/e8532e56943c4db597f816ae3fb6bff4_0 via http://www.acpafl.org/

Table 3 – Datasets from miscellaneous sources

Dataset name	Type of Data	Relevant Fields
Census Tract Populations[12]	American Community Survey's 2012-2016 5-year estimate of people living in Gainesville (57 records, 2012-2016)	Tract ID, Total Population and Demographics
Gainesville Businesses Survey Data[13]	Answers by Gainesville businesses to 44 questions on performance and interactions with the City. (751 records, 2017)	Aggregate measures of business performance and business sentiment
Nationwide Consumer Sentiment[14]	Quantified consumer sentiment regarding the state of the national economy (66 records, 1952-2017)	Index of Consumer Sentiment and Date
Florida Consumer Sentiment[15]	Quantified consumer sentiment of Florida residents regarding the state of the economy (33 records, 1985-2017)	Index of Consumer Sentiment and Date
Gainesville MSA Consumer Sentiment[16] [17]	Consumer sentiment of Gainesville MSA residents on the state of the local economy (25 records, 1993-2017)	Index of Consumer Sentiment and Date
National Gross Domestic Product[18]	Quarterly national Gross Domestic Product measurements by category (88 records, 1929-2016)	Year, Quarter and Real Gross Domestic Product
Gainesville MSA Gross Domestic Product[19]	Yearly city-level Gross Domestic Product measurements by industry for the Alachua and Gilchrest counties (16 records, 2001-2016)	Year, Industry and All Industry Total
Gainesville Per Capita Personal Income[20]	Income distributions of census block groups[21], lists the number of people in income brackets (155 rows, 2016)	Census Tract, Block Group and Total (per ranges of income)

[12] https://factfinder.census.gov/bkmk/table/1.0/en/ACS/16_5YR/DP05/0100000US
[13] Survey data: See Appendix A.
[14] National Consumer Sentiment: http://www.sca.isr.umich.edu/tables.html
[15] Florida Consumer Sentiment: https://www.bebr.ufl.edu/csi-data
[16] The Gainesville Metropolitan Statistical Area includes both Alachua and Gilchrest counties.
[17] Gainesville MSA Consumer Sentiment: provided by the Bureau of Economic and Business Research at UF
[18] https://www.bea.gov/histdata/fileStructDisplay.cfm?HMI=7&DY=2017&DQ=Q3&DV=Third&dNRD=December-22-2017
[19] Gainesville Gross Domestic Product: https://bea.gov/iTable/index_regional.cfm
[20] Gainesville Per Capita Personal Income: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_B19001
[21] Census tracts are composed of several census block groups, which are composed of several census blocks, which are the smallest units of land for which the Census Bureau collects data.

Section II – Gainesville City Businesses

According to the data available on November 1st, 2017, there are approximately 5528[22] active businesses in the City of Gainesville. These businesses can be broadly classified under 20 different categories (following NAICS taxonomy[23]) or more finely categorized (using City of Gainesville's business classes[24]). Figures 1 and 2 show the numbers of businesses of each type[25] for the NAICS and City of Gainesville's categories, respectively. Figure 2 shows only those business categories for which there are more than 50 businesses. The NAICS categorization shows that the most common type of business in Gainesville is retail. 1159 businesses, consisting of 21.97% of all businesses, are classified as retail, followed closely by professional, scientific, and technical services (18.58%, including lawyers, attorneys, tax services, auto repair, laundromats, etc.), and other services (18.13%, including barber shops, beauty parlors, tailors, car cleaning, etc.).

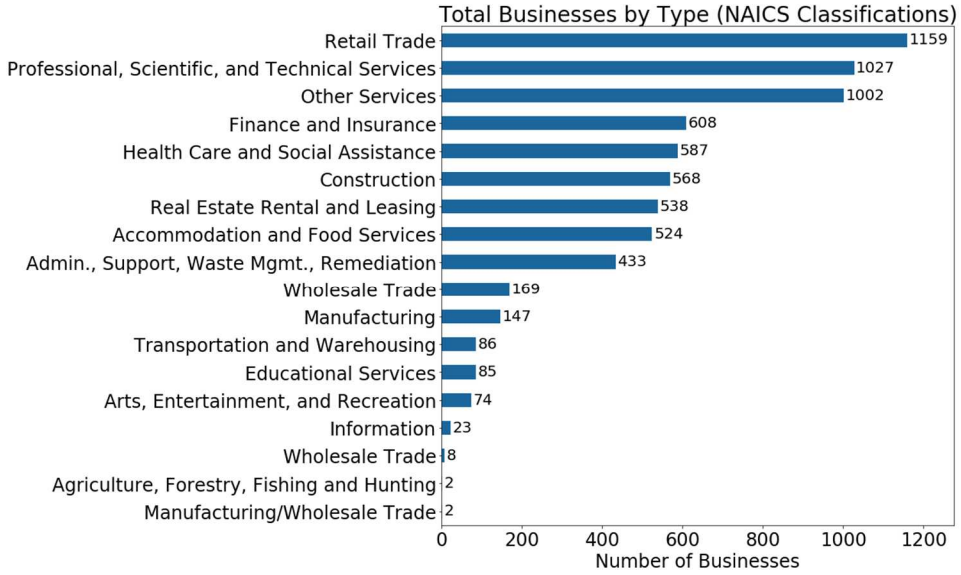

Figure 1 – Total number of active businesses of each type in Gainesville. Businesses are classified according to the NAICS taxonomy.

The life span of currently active businesses ranges from less than one year to just under 70 years. More than half (2938) of these businesses started operating in 2008. In addition to those shown in Figure 3, 225 businesses have been active for

[22] The number of businesses has been inferred from records from different datasets after elimination of records that are likely to be duplicates. Including duplicates would raise the number of active businesses to 7044.
[23] North American Industry Classification System; https://www.census.gov/cgi-bin/sssd/naics/naicsrch?chart=2017
[24] https://data.cityofgainesville.org/Strong-Economy/Active-Businesses/hk2b-em59
[25] Figures 1 and 2 use the full list of 7,044 active businesses; business with multiple types are counted once per category they fall under.

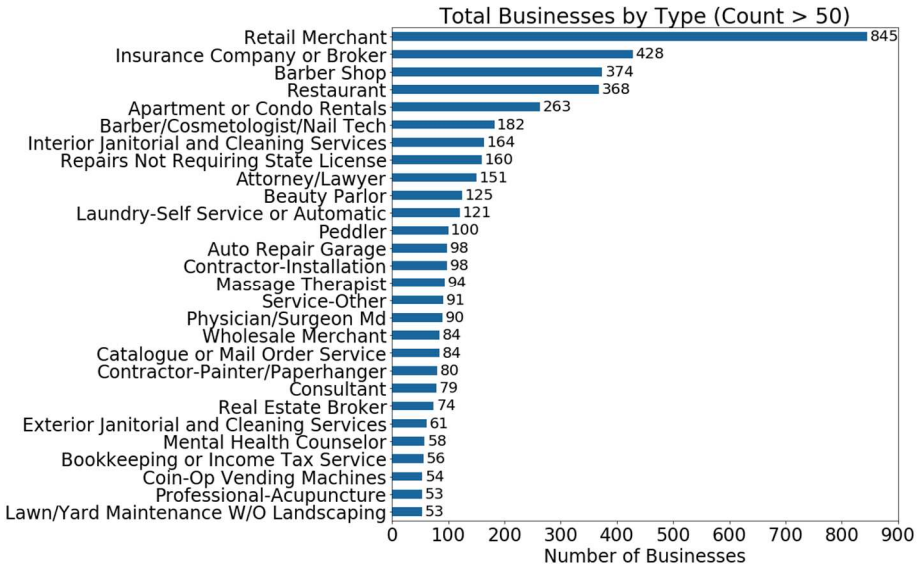

Figure 2 - Types of active businesses in Gainesville for which there are at least 50 instances. Businesses are classified according to the City's taxonomy.

more than 30 years, and only 24 businesses have been in existence for more than 40 years[26].

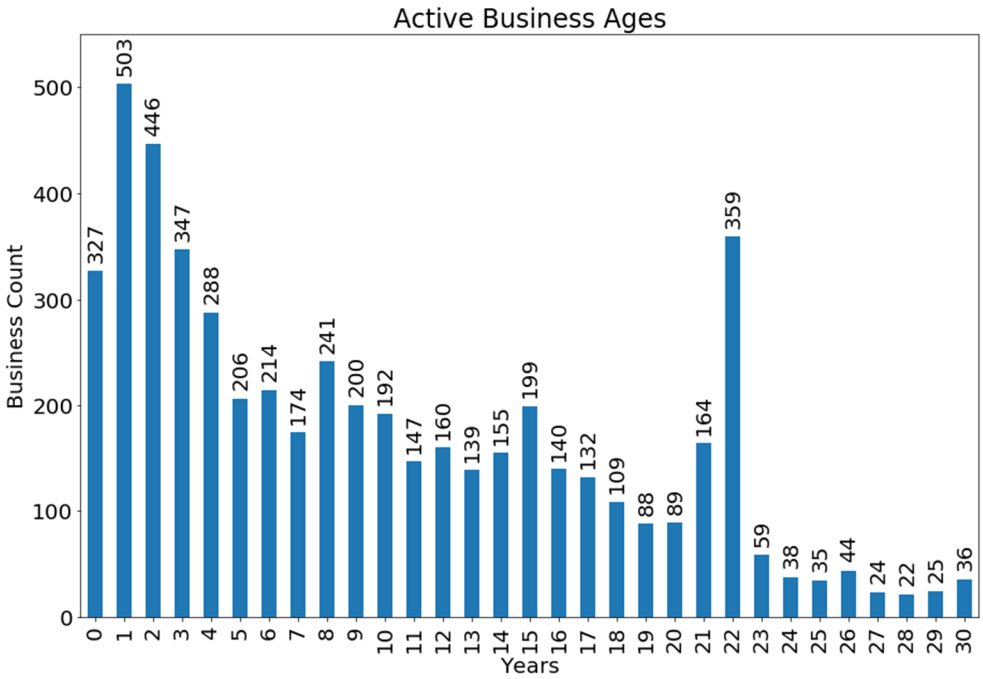

Figure 3 – Bars depict the number of active businesses that have been in business for a given number of years as of 11/01/2017.

The blue bars in Figure 4 show the number of businesses that started in a given year and are still in operation. The orange bars show the number of businesses that started in a given year and are no longer active. From 1999 till 2007, a large number of new businesses were started in Gainesville but a large fraction (76%) of these businesses are no longer active. The number of business closures was the highest

[26] 1995 data (22 years ago in relation to 2017) are inaccurate; record dates reflect record creation dates rather than business start dates; this seems to be an idiosyncrasy of the introduction of new software in the management of data in 1995. This affects only the data for 1995 in several datasets and plots mentioned in this report.

between 2000 and 2008 (between 400 and 800 business closures per year, 560 on average), as shown in Figure 5, which conveys the number of business openings and closures for every year since 1977. Each blue bar in Figure 5 indicates the number of businesses that opened each year and is equivalent to the combined blue and orange bars for the same year in Figure 4. Each orange bar in Figure 5 indicates the number of businesses that closed in a given year.

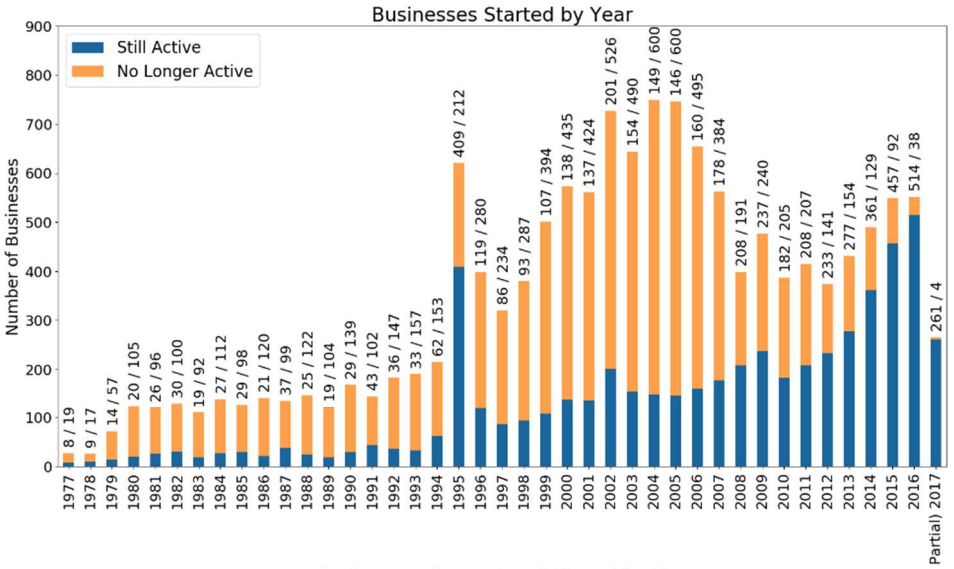

Figure 4 – Blue bars depict the number of businesses that started in a given year and are currently still active. Orange bars depict the number of businesses that started in a given year but are no longer active.

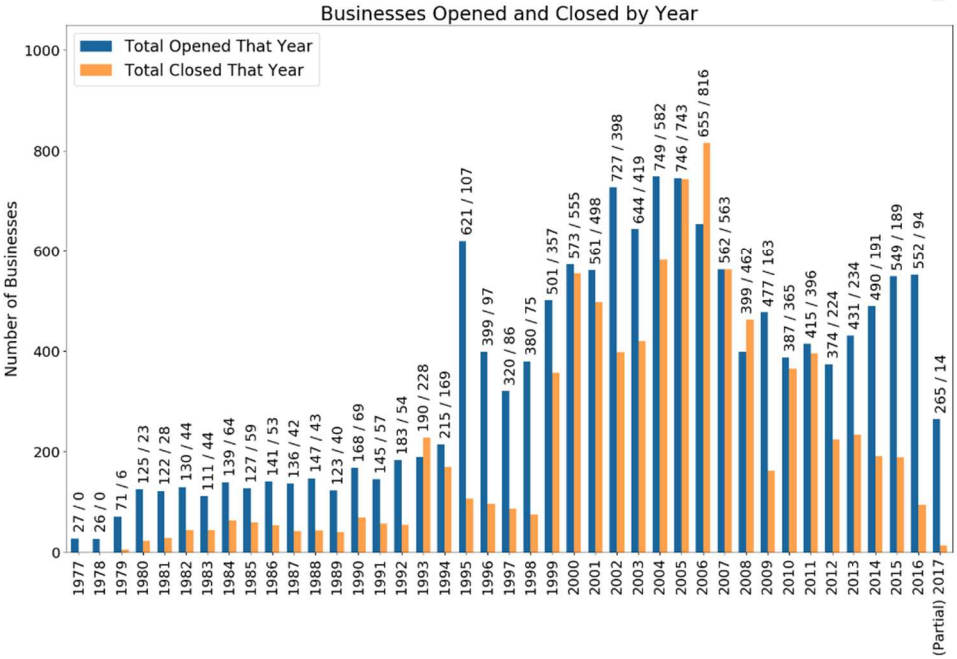

Figure 5 - Blue bars depict the number of businesses that started in a given year. Orange bars depict the number of businesses that ceased operation in a given year.

Out of 8,651 total inactive businesses (i.e. businesses that started since 1950 but have ceased to exist), only 1,272 have their business type listed. Figure 6 shows the number of business closures (i.e. businesses that are inactive) for different types of businesses. Types of businesses for which there were fewer than six closures are not shown. Event-oriented businesses (i.e. 378 peddlers and 174 yard-parking services) have been excluded from Figure 6 since they are transient in nature).

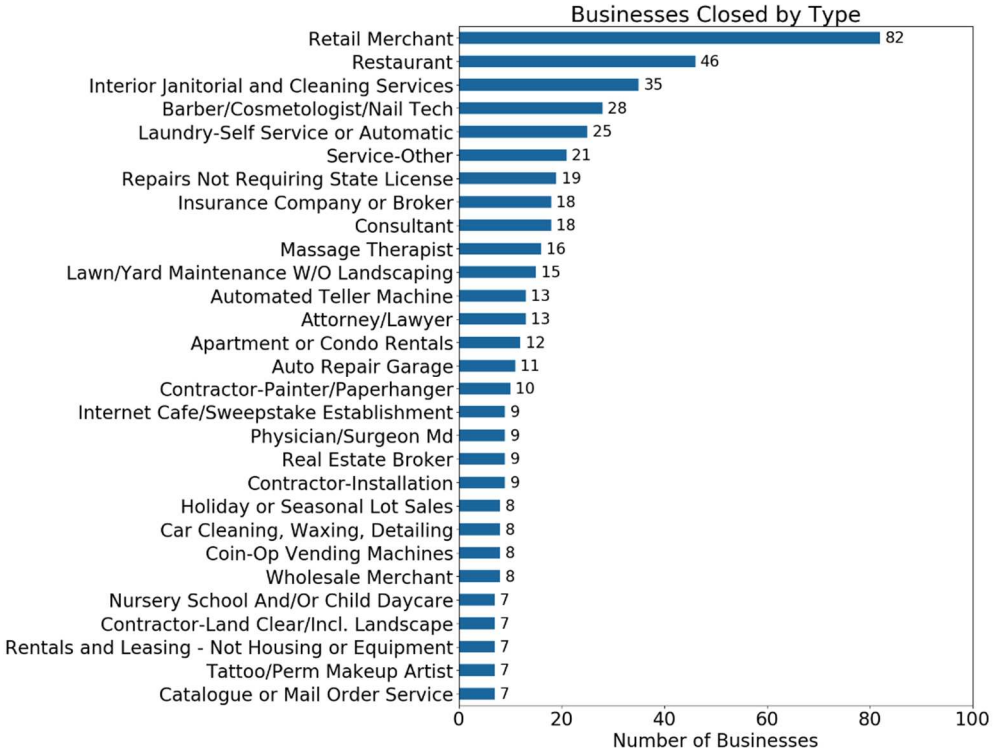

Figure 6 - Number of businesses of a given type that have closed since 1950. Business types are in accordance with the City's taxonomy. Business types with fewer than 5 inactive businesses are not shown. Transient businesses (peddlers and yard-parking services) also are not shown.

Geographical distribution of Gainesville businesses

The geographical distribution[27] of the 4,972 active business locations in Gainesville shows that 60% of them (2,968) are located in the NW part of the city (see Figure 7). The SE part of Gainesville is home to only 283 active businesses which corresponds to approximately 6% of the total. In comparison, the population distribution[28] shows that 41% of the Gainesville residents live in the NW while just 3%[29] live in the SE. The population-to-business ratios of the NW and NE quadrants are approximately 17 and 15 people per business, respectively, whereas the SW is far more residential (54 people per business) and the SE is only slightly more residential (21 people per business). The SW is the only region that contains a proportion of Gainesville's population (43%) that is larger than its proportion of Gainesville's businesses (20%).

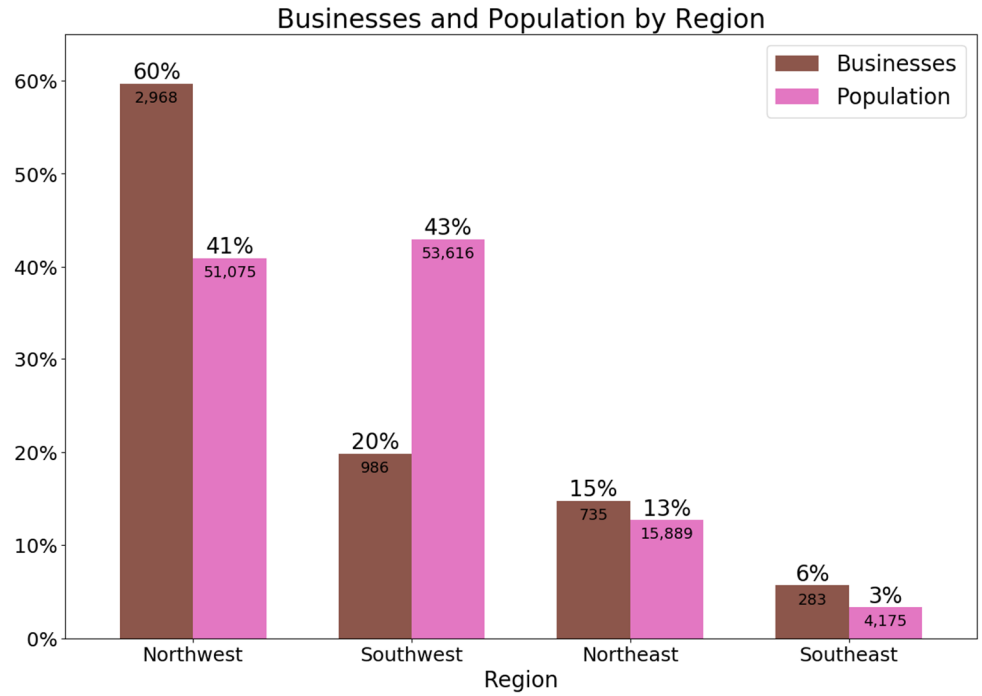

Figure 7 - Distribution of currently active businesses and population across the NW, SW, NE and SE regions of Gainesville.

The distribution of active businesses in Gainesville can also be viewed according to census tracts shown in Figure 8 as labeled by the United States Census Bureau. Figure 9 shows that the downtown tracts of Gainesville have the highest numbers of active businesses. Figure 10 shows a high population concentration around the

[27] Gainesville was divided into four geographic quadrants defined by the intersection between Main Street and University. Businesses with multiple entries in the dataset were only counted once per unique location. Businesses located that are either unspecified or outside of Gainesville are not included in the count.

[28] Population data was taken from census tract populations provided in the American Community Survey's 2012-2016 five-year estimate. For census tracts that are only partially inside Gainesville's municipal boundaries, the population was estimated according to the fraction of the tract's land that lies within those boundaries.

[29] This number may be low due to the methods we used to estimate population. See above footnote.

University of Florida, which contributes to the SW's relatively high population compared to the other regions. Figure 11 shows that tracts in the SW neighborhood of the University of Florida have the highest number of residents per business. The identification of possible causes for the current geographical distribution of businesses is out of this project's scope.

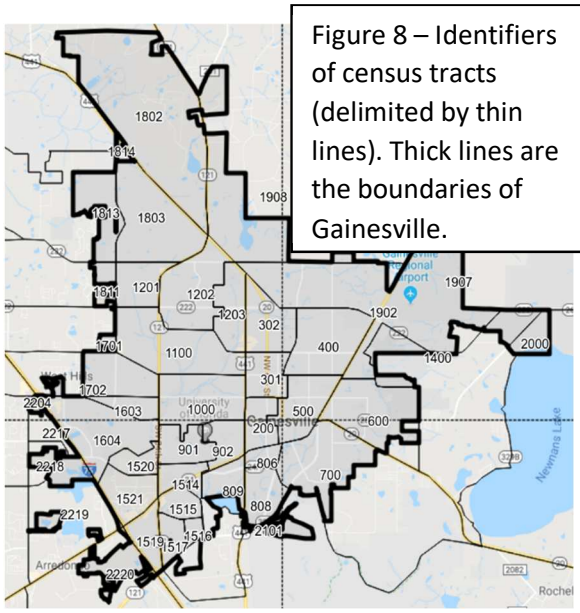

Figure 8 – Identifiers of census tracts (delimited by thin lines). Thick lines are the boundaries of Gainesville.

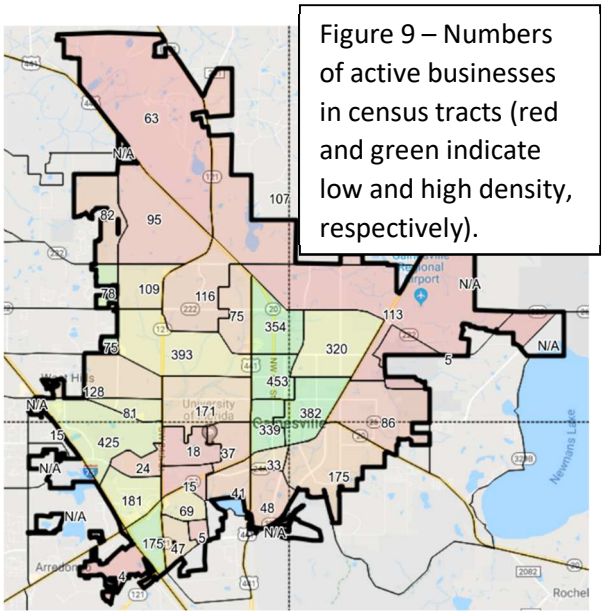

Figure 9 – Numbers of active businesses in census tracts (red and green indicate low and high density, respectively).

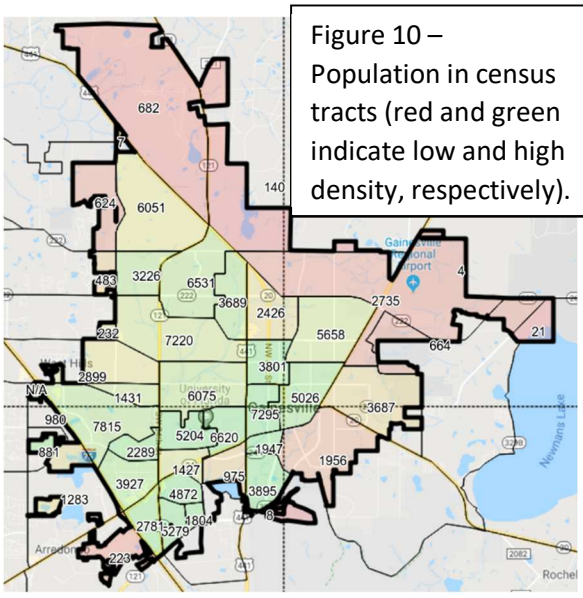

Figure 10 – Population in census tracts (red and green indicate low and high density, respectively).

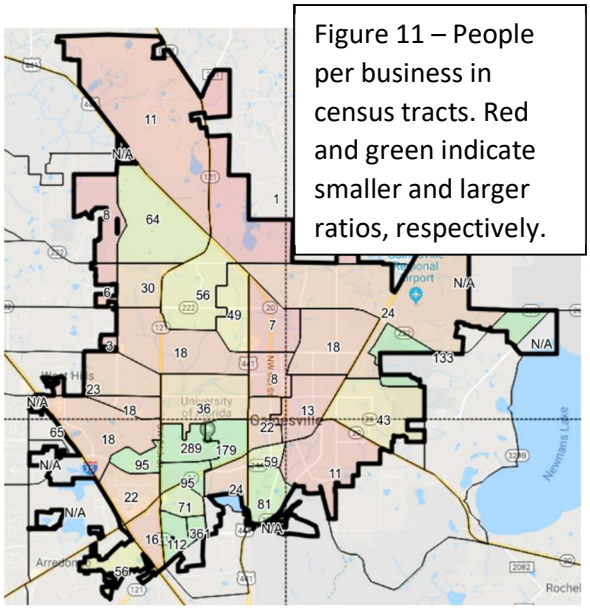

Figure 11 – People per business in census tracts. Red and green indicate smaller and larger ratios, respectively.

Section III – Gainesville City Businesses Sentiment

We conducted a survey[30] of the businesses in Gainesville to gauge their sentiment of how successful they are and their satisfaction with their interactions with the City. Of those that started the survey, 479 completed it, 165 partially completed it. The response rate was only 16.9%, which is typical. The results of the survey are intended to give us information about all businesses located within the city limits of Gainesville. The conclusions from this survey are based on this smaller subset of Gainesville businesses[31] whose geographical distribution is shown in Figure 12.

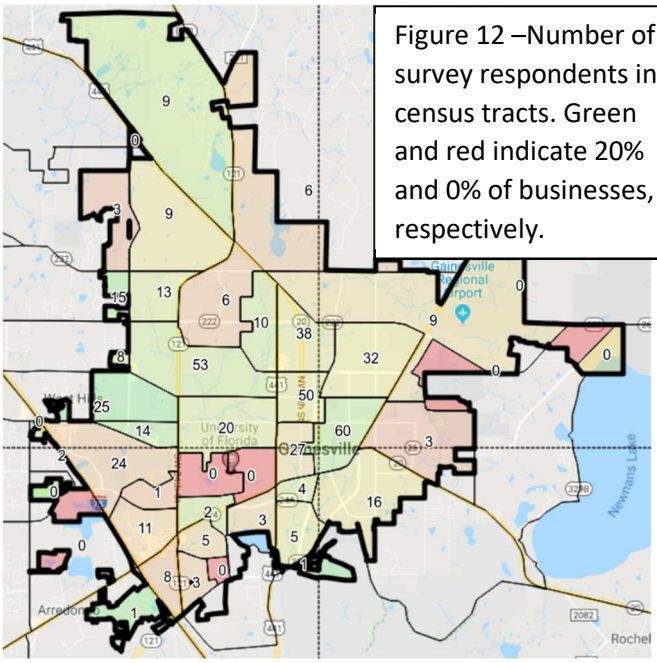

Figure 12 –Number of survey respondents in census tracts. Green and red indicate 20% and 0% of businesses, respectively.

The large majority (more than 95%) of the respondent businesses have fewer than 50 employees. The reported start dates of these businesses show the same trend as the blue bars of Figure 3 -- more than 65% and 40% started since 2001 and 2011, respectively. When asked about the success of their business, 45.82% answered "very successful" while 46.2% answered "somewhat successful."

The survey asked whether businesses would like the City to provide services during each of the phases of their business life mentioned in the introduction of this report. For each of the thirteen questions, at least 30% of those who responded indicated that they would be interested in receiving City services. The "legalize" and "finance" phases had the most interest (47% and 45%, respectively) and the "hire" and "taxes" phases had the least interest (both at 32%). Of the 64 respondents who said they made use of

[30] The list of businesses was obtained from the Active Businesses list available on the City of Gainesville website, with those indicating location outside the city limits removed. This list was then de-duplicated by business and email address resulting in 4,000 business's email addresses. On October 17, 2017, an email invitation to take the survey was sent to these 4,000 businesses. Reminder emails were sent on October 25, November 2, and November 8. The survey closed on November 13. 179 emails bounced and 97 respondents opted out of taking the survey. 4 respondents were ineligible. A business was eligible to take the survey if they were open at the time of the survey or their business closed in 2016 or 2017.

[31] The survey results are weighted by the business type found in the Active Businesses list. Weighting the data by business type is intended to correct for lack of representativeness of the survey respondents, reflecting a refinement in the percentages reported.

services from the City of Gainesville in the last 12 months, 44 indicated that they were able to find the services they were seeking from the City and 37 want to see improvement in these City services

The survey also asked each business about its sentiment regarding interactions they might have had with the City when searching for a business location and/or drawing development or building plans. The survey respondents were asked to rate on the scale of 0 (worst) to 10 (best) the service received from several City Departments, namely the Department of Doing (Planning and Building Construction), the Public Works Department, Gainesville Regional Utilities, the Budget and Finance Department and the Community Redevelopment Agency. All departments had an average rating higher than 6 (see Figure 13).

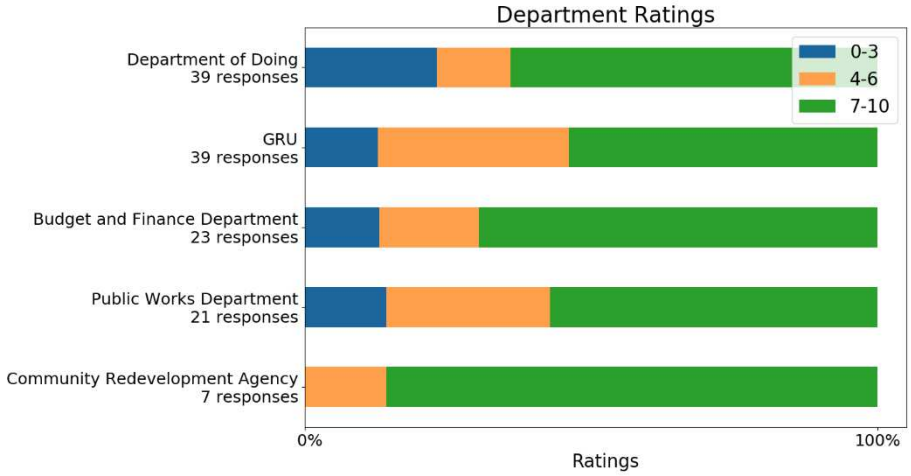

Figure 13 – Distribution of ratings by survey respondents of services provided by City departments. Lengths of bars represent the percentage of respondents who selected a number in that range.

The survey asked businesses to state to what degree they are likely to recommend opening a business in Gainesville. Approximately 43% of the survey respondents stated that they are extremely likely or very likely to do so whereas approximately 23% indicated that they are slightly or not at all likely to do it. Approximately 47% of the businesses who answered this question indicated that their answer was not influenced by their experience with the City of Gainesville.

The businesses who reported being either “very successful” or “somewhat successful” attributed their success largely to their own efforts and personal experience in the field. The 8% of business owners who reported being either “somewhat unsuccessful” or “very unsuccessful” considered their lack of success the result of circumstantial reasons (47%) and, to a smaller extent, of their interactions with the city (22%). Regardless of how successful respondents perceived their businesses, about a third of the respondents reported that the state of the local economy affected their success either positively or negatively.

Section IV – Combined Data and Survey Results Analysis

The survey data and the datasets described in previous Sections I and II were combined in an attempt to find out whether there was a meaningful correlation between publicly available data and the likelihood of business success. The following factors were considered:

- Gross domestic product (GDP) and GDP growth rate
- Consumer sentiment index (CSI)
- Business age
- Electricity consumption
- Crime numbers
- Building permits

There is a strong correlation (0.90) between the GDP of the US and the GDP of the Gainesville Metropolitan Statistical Area (which consists of both Alachua County and Gilchrest County). Figure 14 shows that there is a high correlation (about 0.82) between national GDP and the total number of active businesses in Gainesville.

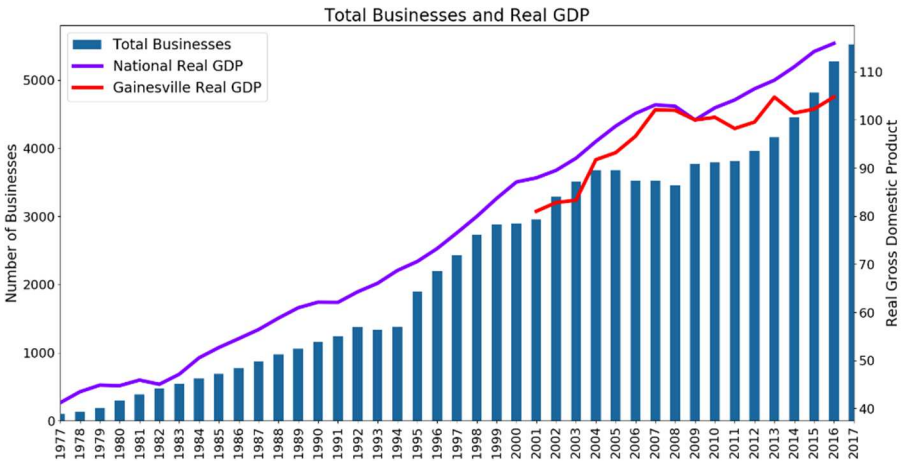

Figure 14 – National and Gainesville MSA gross domestic products (purple and red lines, respectively) alongside the total number of businesses in Gainesville for each year.

When considering the yearly net change in the total number of businesses versus GDP in Figure 15, the correlation is much less apparent. This discrepancy is especially due to the relatively unstable nature of the local GDP (compared to the national) which lacks enough data to establish any clear trends. Figure 15 considers only the years after 1996 since no data is available for GDP of Gainesville before 2001 and Gainesville business data prior to 1997 is not entirely reliable due to reported changes in the information technology used for tracking business activity.

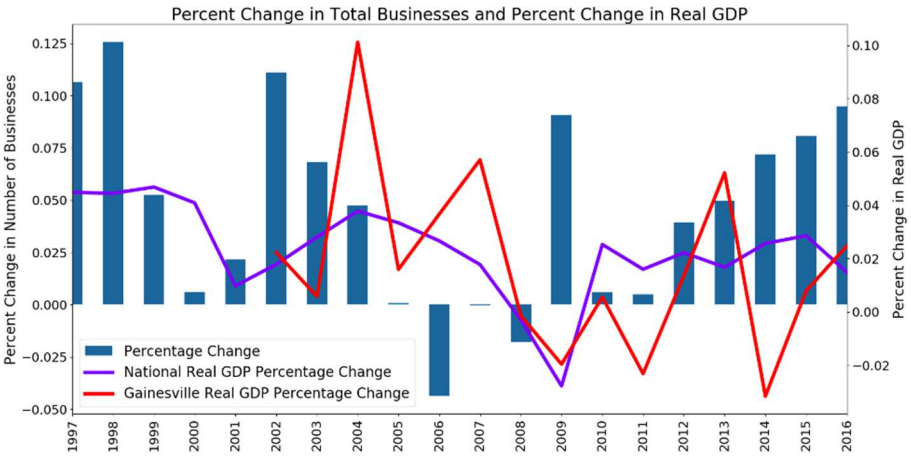

Figure 15 – Change in National and Gainesville gross domestic products (purple and red lines, respectively) and the percent change in the total number of businesses in Gainesville for each year.

Similarly to GDP, there is a high correlation (0.97) between consumer sentiments at the local and national levels. The correlation of consumer sentiment and business creation in Gainesville in the same year is low. However, the correlation of national consumer sentiment and business creation two years earlier is high (0.72) (this corresponds to shifting the sentiment lines in Figure 16 two years to the left), which suggests there may be delay between actual business growth and public perception.

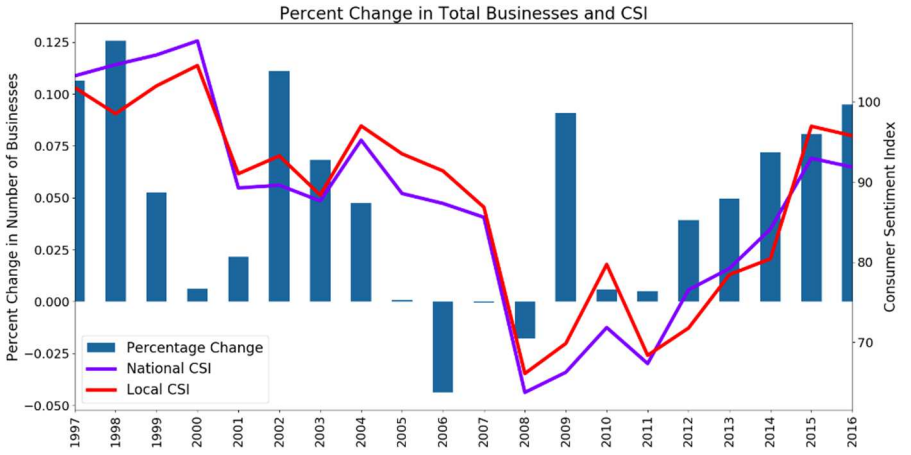

Figure 16 –National and Gainesville MSA consumer sentiment indexes (purple and red lines, respectively) alongside the percent change in the total number of businesses (bars) each year.

The correlation factor between business age and success is relatively high (0.70). In order to find this correlation, for each business that answered the survey questions, its lifespan was calculated and correlated with their answer to the question on whether they considered themselves successful. The collected results for 519 respondents are shown in Figure 17. The represented businesses' ages range from 0 to 42 years. The first decade of a business's lifecycle seems to be the most uncertain, while businesses that have survived for well over a decade tend to maintain their success. Of the 155 respondents whose businesses were at least 16 years old, only 1 reported being somewhat unsuccessful.

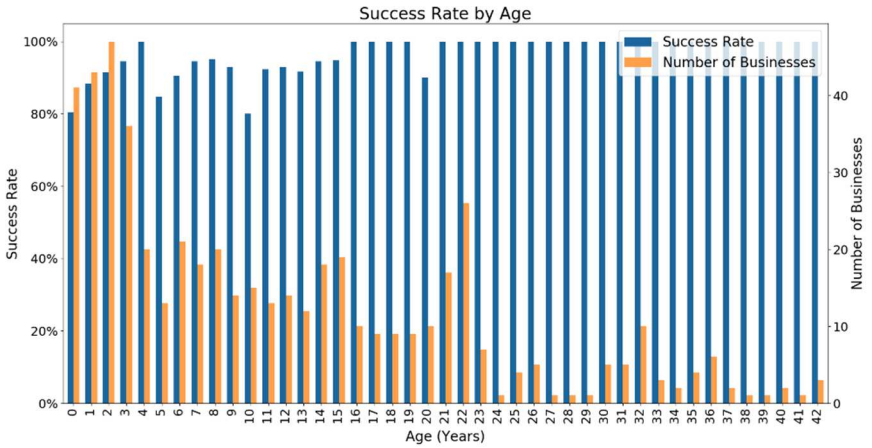

Figure 17 - Number of businesses of a given age that filled out the survey (in orange) and their success rate (in blue).

The conclusion that there is a high correlation between business age and success makes sense but should still be critically considered. With 92% of survey respondents reporting being successful, we know what successful businesses look like, but with only 8% reporting being unsuccessful, we don't really know what unsuccessful ones look like (in terms of data). This argues the case for a similar survey being given annually, leading to eventually having enough data to profile unsuccessful businesses. Some businesses that reported being successful are bound to close down, and we can see what data trends occurred in the months leading up to their failure. Also, from Figure 5, very few businesses are actually failing in recent years. One also wonders whether there is an age beyond which business performance is likely to decrease and lead to closure. We need data on predictors of business closures during periods of bad economic performance. It would have helped to have data for the years in which businesses failed. Unfortunately, 2017 data do not help determine why a business failed in 2005.

We found no correlation between each of the other five factors (electricity consumption, crime numbers, building permits, building code violations, and zoning violations) and business success (correlation values were all below 0.1).

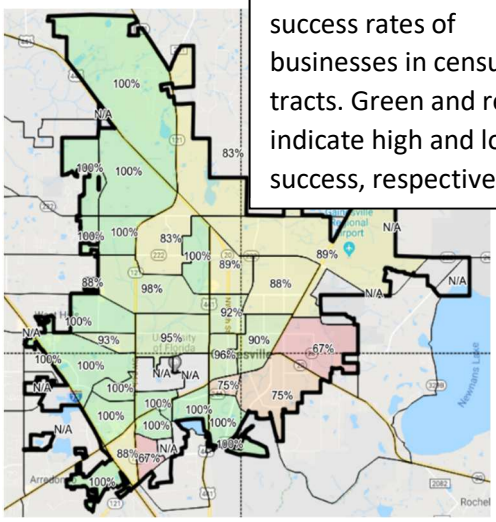

Figure 18 – Average success rates of businesses in census tracts. Green and red indicate high and low success, respectively.

The business success rates change across Gainesville as shown in Figure 18, where all but four of the forty-four census tracts have business success rates above 80%. Three occur around the southeast quadrant of Gainesville and one in the southern extent of the southwest region.

Section V – Data Management Practices and Analysis Tools

Data availability and accessibility

The City of Gainesville Open Data portal provides convenient access to a variety of datasets, including those mentioned in Table 1. The portal is a potentially useful resource for data-centric studies of Gainesville businesses – it should be continued and expanded. However, currently these data are insufficient or inadequate to conduct meaningful analysis. The data identified in this study (see Table 2) helped in enabling limited modeling as discussed in Section 4. These data, the survey data and additional data should be collected and made more easily accessible in the future. Also needed are more detailed documentation of what each dataset contains, how data are formatted, whether and what vocabularies are used and their logic organization schema, which would help in efficiently and effectively using the data.

To enable more sophisticated analysis, data are needed beyond those available for this study. It is clear that the availability of more data for everyday business operations and their economic performance will increase the likelihood of developing good predictive data-centric models. However, laws and regulations that protect private business and personal data will make it hard to have direct access to explicit information on individual business performance. A survey such as the one used in this study can collect some of the desired data, but the survey should be brief enough that businesspeople will be willing to participate. The prospect of indirectly getting business performance information from social networks, web sites and other media also deserves further investigation. However, recent regulations from governments and social media companies might limit access to those data to less than a useful level.

Data quality

The quality and completeness of the data available for this study need to be improved if similar data are to be useful for future studies. In particular, numerous records have erroneous entries that cannot be automatically corrected thus requiring human intervention that can be quite laborious. Examples of errors, omissions and idiosyncrasies found in datasets used in this study are shown below in Table 4 and further discussed in the ensuing text.

Table 4 – Types and examples of poor data-quality problems in existing datasets

Problem	Example	Affected Datasets
No shared identifier for businesses across datasets	*Active Businesses, Inactive Business* use an "ID"; *Building Code Violations, Zoning Violations, GRU Electric Data* use "parcel"	*All*
Multiple entries for one business	21.5% of businesses in *Active Businesses* have several entries in the dataset; 1.7% of inactive businesses have multiple entries	*Active Businesses* *Inactive Businesses*
Mislabeled duplicate business names	Businesses with the same name have an arbitrary number of periods appended to their names to distinguish them. This makes their names inconsistent with other datasets. e.g. "THE PLACE." and "THE PLACE.."	*Inactive Businesses*
Inconsistent address formats	Both SUITE and STE are used interchangeably. Other examples include NORTH versus N and STREET versus ST. Some addresses are missing their road type ("Rd", "St", etc.)	*Active Businesses* *Inactive Businesses* *Parcel Data*
Inconsistent name formats	"THE" placement, e.g. "PLACE, THE" vs. "THE PLACE"	*Active Businesses* *Inactive Businesses*
Inconsistent date formats	*Active Businesses, Inactive Businesses* use YYYY-MM-DD; *Building Permits* uses MM/DD/YYYY; *GRU Electric Data* uses YYYYMMDD and YYYY/MM	*All datasets*
Inconsistent parcel formats	*Building Permits*, *GRU Electric Data* use XXXXX-XXX-XXX, *Zoning Violations* and *Code Violations* use XXXXXXXXXXX	*Building Permits* *GRU Electric Data* *Zoning Violations* *Code Violations*
Typos in shared labels	"Acccessary Structure" vs. "Accessory Structure"	*Building Permits*
Incorrect latitude/longitude	Coordinates under "Location" in *Active Businesses* used for the same street name in other cities; e.g. "4010 Newberry Rd" was given coordinates in Winnsboro, SC	*Active Businesses*
Missing business types	7,379 inactive businesses (85%) do not have a business type listed	*Inactive Businesses*
Missing latitude/longitude	203 active businesses (3%) with valid addresses are missing latitude/ longitude coordinates under "location"	*Active Businesses*
Erroneous dates	713 inactive businesses (8%) have "out of business" dates that are earlier than their "start dates"	*Inactive Businesses*
Filler data	Business addresses listed as "VARIOUS" have the same arbitrary "location" coordinates (25.810289, -80.297266)*;* Fake phone numbers, e.g. 63 businesses have 352-123-4567 listed	*Active Businesses* *Inactive Businesses*

Data-entry validation mechanisms that check for proper formatting of data and compliance with allowed vocabularies could help mitigate some of the errors found in the data. Checking against other datasets would also be helpful in detecting wrong data that are provided negligently or maliciously. For example, a fake phone number was often provided as 352-123-4567 which, while following the right format, would fail a verification test that looks up the name of the owner of the phone number or sends a message to the owner for confirmation purposes. Similarly, area codes, street addresses and dates are often partially verifiable at data-entry time. In the absence of data-entry validation and best practices, only a limited amount of data cleaning might be possible in automated ways in a few cases (examples include the conversion of formats of dates or addresses and the correction of typos).

A major challenge faced by this study is the sparseness of the data. Very few businesses have data in all the different datasets. Thus, it is difficult to give a clear binary answer about the success of a business from the existing data. Another example of incomplete data is the absence of business-type information in most datasets, further limiting any analysis that takes into consideration the type of business.

Another challenge is the presence of duplicate entries in a single dataset. For example, the Active Businesses dataset has multiple entries that differ only in business type. It is hard to automatically correct such entries because they could either be redundant (thus candidates for deletion) or legitimate entries that describe individual entities that are engaged in multiple kinds of businesses.

The use of taxonomies or limited vocabularies is also desirable. For example, business types could be consistently described using NAICS codes. This would facilitate data analysis per business type as well as correlation with other datasets or studies that might use the NAICS classification.

If possible, the Inactive Businesses dataset needs to be maintained alongside the Active Businesses dataset to have accurate and more complete view of creation and closure of businesses (Currently the inactive businesses dataset misses critical information such as the business type data for the majority of records, as well as erroneous data such as negative lifespans).

Data organization

Desirably, different organizations in the city should join their efforts to come up with compatible data schemas that share at least one key to uniquely identify business entities. This would reduce the efforts required in data gathering and cleaning and allow researchers to devote more time analyzing the data. Data formats need to be agreed upon or, at least, well documented and enforced for each dataset. This applies, for example, to geographical addresses across different datasets so they can be efficiently and accurately matched using software. This will help greatly in merging and analyzing multiple datasets.

Data analytics

Data-centric derivation of models -- predictive or retrospective -- requires the querying of multiple datasets for multiple occurrences of specific data values (or ranges of values) and data relationships. In many cases, answers to complex queries require computational power to examine large amounts of data and to compute "what if" scenarios that test hypothesized models. Additionally, where users' validation or intuition need to be supported, visualization of the results of queries and data analytics are desirable.

In this study we used standard tools and technologies for data analytics and visualization, relying on open-source database and spreadsheet software[32] to merge and query data originally available as spreadsheets. This approach was adequate for the limited number of datasets of rather small sizes considered in this study. As simple examples of queries, the total number of businesses was found by counting the number of businesses with a unique "name" and "address", whereas querying for unique "name" and "business type" produced the number of businesses for each type. For purposes of visualization, we also used open-source tools to generate the graphs and maps[33], as exemplified by the figures included in this report.

Repeatability and reproducibility

The tools used in this study were chosen to so that all results are reproducible using the same data. All datasets were downloaded as individual spreadsheets and only

[32] All data analysis and graphing were completed using the tools provided in the Anaconda Distribution (https://www.anaconda.com/), including Jupyter Notebook (http://jupyter.org/) running Python 2.7.

[33] Map visualizations were completed using the Google Maps JavaScript API (https://cloud.google.com/maps-platform/) running on a Node.js server (https://nodejs.org/en/). Data was housed using MySQL.

data dated prior to November 1st, 2017 were used. Refer to Table 1 for a detailed list of the datasets used and where they were accessed.

A comprehensive analysis of the available data required the composition of a single, unified dataset. The datasets had to be manually manipulated to make unification possible. Datasets with different business identifiers were mapped to each other in order to identify the businesses being described. This was done by assigning a unique ID to each business, which were then matched to the extent possible by using name, parcel number (where available) and GRU's internal ID provided in the Active Businesses and Inactive Businesses datasets. A cross analysis of the available data required the shared data fields to be renamed, reformatted, and normalized. This included names, dates, addresses, numerical identifiers, and timespans.

Future studies using newly available data will require the above-described manual efforts to be done anew. However, the automated parts of this study can be reproduced from documented analysis scripts[34].

Scalability

The size and number of datasets for the kinds of analysis needed for this and similar studies are not very large. As long as future datasets contain a negligible number of errors that need to be manually corrected, the methods used in this analysis are very general and can scale to any amount of data. Analysis was automated on entire datasets, not done manually.

It is entirely possible and desirable to investigate new types of data whose size might call for advanced data analytics capable of processing huge amounts of data. Open-source analytics tools and locally available expertise are already available to handle such scenarios. However, a discussion of what such data might be and suitable approaches to analyze them is outside the scope of this study.

[34] Access to these scripts, as well as a snapshot of the publicly available data used to conduct the analysis, is available online at https://github.com/acislab/DGLIM

Section VI – Conclusions

The aim of this project was to enable the use of heterogeneous data sources and IT tools to identify, quantify, characterize and correlate key factors for success and failure during the lifecycle of different types of businesses in Gainesville. We conjectured that data collected for City administrative purposes and data provided by government agencies and other open sources could be used to study characteristics of Gainesville businesses.

We were able to derive interesting observations from the data we collected, including the different types of Gainesville businesses and their distribution, chronological trends of their numbers, annual rates of business creation and closure, geographical distributions of all businesses and correlations between Gainesville business performance and several economic indicators, including local and national gross domestic products and consumer sentiment indexes. We also identified factors for which we could not find any relation to business performance.

The purpose of our study was to demonstrate the viability and potential value of bringing together and analyzing available Gainesville-related datasets. It was not our objective to conduct either a study of the economic performance of Gainesville or its business environment. Consequently, we refrained from trying to claim too much from the data as far as performance and impact of Gainesville businesses, since we also realize that such claims require deeper inquiries beyond looking only at the available data. For example, business starts may be due to either a good economy that creates business opportunities or a bad economy that requires unemployed individuals to become self-employed – further analysis is needed to differentiate these cases. Such analysis could take into consideration unemployment data which, as an example, shows that 2009 was a year of high unemployment in Gainesville coinciding with a spike in business creation as depicted in Figures 14 and 15. Additional analyses are also needed to take into consideration county economics, transportation corridors (e.g. I-75) and other factors beyond Gainesville city limits. Nevertheless, the above-mentioned observations and the methods used to reach them could provide information of potential value for such studies from which additional questions could be answered or quantitative evidence could be derived to support qualitative claims.

This study was possible only because Gainesville is one of the leading cities in the USA using an open data approach that provides to the public access to a large

number of Gainesville-related data. However, this study also revealed limitations and challenges of data-centric analytics using available data. Significant problems exist regarding the quality, completeness, uniformity, unique identification, representation and logical organization of data from multiple datasets maintained by the City of Gainesville and associated entities. These problems limit the extent, generalizability and reliability of studies using available data.

Our main conclusion is the confirmation that there are significant benefits and opportunities to be gained from a City-level data management and analytics framework that enables the collection of high-quality data along with logic schemas that enable cross-dataset queries. While this project's focus was on data that might be relevant to Gainesville businesses, a City framework for data management and analytics should be general enough to also support other types of City-relevant data. For example, such data might result from environmental sensing, traffic monitoring, public transportation monitors, citizen-provided information, personal devices, City-deployed sensors, operational records from private and public entities who wish to share data, etc.

Other cities in the US and elsewhere are exploring how such a framework should look like and be deployed. There is also growing interest from industry and businesses in providing solutions for the whole or parts of the framework. The City of Gainesville should leverage these ongoing developments and explore how they might suit the needs of Gainesville as it becomes a "New American City".

In closing, this project brought together researchers from different departments at the University of Florida and the City of Gainesville. The team met, on average, once every two weeks for approximately nine months and reached out to several branches of the City government and Gainesville businesses. Beyond this report, the joint work contributed inputs to activities of the Department of Doing and other departments of the City of Gainesville. It also provided educational experiences to two graduate students at UF. Continuing encouragement and funding of interactions between UF and City researchers would further contribute to improved City operations and opportunities for education and research at UF that address challenges to cities such as Gainesville.

Appendix A – 2018 Gainesville Business Survey Data Summary

On October 17, 2017, an email invitation to take the survey was sent to 4,000 businesses in Gainesville. Reminder emails were sent on October 25, November 2, and November 8. The survey closed on November 13. 179 emails bounced and 97 respondents opted out of taking the survey. A business was eligible to take the survey if they were open at the time of the survey or their business closed in 2016 or 2017. Of those that started the survey, 479 completed it, 165 partially completed it, and 4 were ineligible. The response rate is 16.9%.

Survey questions and response statistics

The table below lists the survey questions and provides statistical summaries for the responses. The column labeled N shows counts for each possible answer, while % shows the corresponding percentages (these are non-weighted values). The column labeled W_N shows weighted answer counts, while W% shows the corresponding percentages. The weighted values are used to improve how well the survey data analysis represents the different types of businesses in the City of Gainesville. The business types considered for the survey are in accordance with the Active Businesses List maintained by the City on its Open Data portal. The presence of an asterisk (*) indicates a multiple-selection question where percentages may total more than 100%.

Q1	**What year did this location of your business start in Gainesville? If your business moved from another location to this location, please answer with the year your business moved to the new location.**				
	Answer	**N**	**%**	**W_N**	**W%**
	Before 1960	13	1.96%	9.81	1.65%
	1961 - 1970	8	1.20%	6.82	1.14%
	1971 - 1980	42	6.33%	32.20	5.40%
	1981 - 1990	61	9.19%	59.16	9.94%
	1991 - 2000	83	12.50%	81.59	13.73%
	2001 - 2010	170	25.60%	158.59	26.67%
	2011 - 2017	287	43.22%	246.48	41.47%
	Total	664	100%	594.65	100.00%
Q2	**Is this location of your business still open or has this location closed?** *(Skip to question q4 if selected Open)*				
	Answer	**N**	**%**	**W_N**	**W%**
	Open	635	95.20%	564.73	95.11%
	Closed	32	4.80%	29.03	4.89%
	Total	667	100%	593.76	100.00%

Q3	**In what year did this location of your business close?** *(Asked if the answer given for Q2 was "Closed")*				
	Answer	**N**	**%**	**W_N**	**W%**
	2008	1	3.45%	0.57	2.28%
	2015	3	10.34%	2.34	9.38%
	2016	4	13.79%	5.44	21.79%
	2017	21	72.41%	16.61	66.55%
	Total	29	100%	24.96	100.00%
Q4	**How many employees currently work at this location of your business? Please include all full and part time employees. Do not include anyone that did work as an independent contractor or subcontractor.**				
	Answer	**N**	**%**	**W_N**	**W%**
	0	54	8.96%	55.71	10.50%
	1 - 10	412	68.33%	343.80	64.82%
	11 - 20	69	11.44%	61.01	11.50%
	21 - 30	25	4.15%	28.69	5.40%
	31 - 50	20	3.32%	19.82	3.74%
	51 - 100	19	3.15%	17.35	3.27%
	101 - 200	3	0.49%	3.00	0.58%
	201 or more	1	0.17%	1.02	0.19%
	Total	603	100%	530.40	100.00%
Q5	**How many employees worked at this location of your business as of January 1, 2017? Please include all full and part time employees. Do not include anyone that did work as an independent contractor or subcontractor**				
	Answer	**N**	**%**	**W_N**	**W%**
	0	70	11.69%	67.80	12.89%
	1 - 10	398	66.44%	332.59	63.21%
	11 - 20	72	12.02%	67.13	12.76%
	21 - 30	23	3.84%	24.76	4.71%
	31 - 50	14	2.34%	13.38	2.54%
	51 - 100	18	3.01%	16.48	3.13%
	101 - 200	3	0.49%	3.00	0.57%
	201 or more	1	0.17%	1.02	0.19%
	Total	599	100%	526.16	100.00%
Q6	**How many employees worked at this location of your business when it closed? Please include all full and part time employees. Do not include anyone that did work as an independent contractor or subcontractor**				
	Answer	**N**	**%**	**W_N**	**W%**
	1	11	45.83%	9.59	44.40%
	2	2	8.33%	1.08	5.00%
	3	2	8.33%	1.83	8.47%
	4	1	4.17%	1.06	4.91%
	5	2	8.33%	2.04	9.44%
	6	2	8.33%	2.33	10.79%
	7	1	4.17%	1.06	4.91%
	10	1	4.17%	0.46	2.13%
	15	1	4.17%	1.27	5.88%
	27	1	4.17%	0.88	4.07%
	Total	24	100%	21.6	100.00%

The next questions are about the phases a business might go through.

Q7	**The dream phase involves activities such as brainstorming ideas for a business. Would you have liked the City of Gainesville to provide services when your business was in the dream phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	215	35.13%	188.9	35.22%
	No	397	64.87%	347.44	64.78%
	Total	612	100%	536.34	100.00%
Q8	**During the plan phase, a business person will think through how the business will run and draw up a business plan. Would you like the City of Gainesville to provide services during the plan phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	234	38.94%	200.38	37.98%
	No	367	61.06%	327.25	62.02%
	Total	601	100%	527.63	100.00%
Q9	**During the finance phase, money is obtained by pooling savings or gifts, or obtaining loans or venture capital. Would you like the City of Gainesville to provide services during the finance phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	266	44.56%	235.91	44.99%
	No	331	55.44%	288.46	55.01%
	Total	597	100%	524.37	100.00%
Q10	**During the legalize phase, the business undertakes activities such as incorporating or establishing the business as a sole proprietorship. Would you like the City of Gainesville to provide services during the legalize phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	279	47.05%	242.55	46.52%
	No	314	52.95%	278.72	53.48%
	Total	593	100%	521.27	100.00%
Q11	**During the brand phase, the business develops and implements a marketing plan. Would you like the City of Gainesville to provide services during the brand phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	223	38.18%	190.44	37.12%
	No	361	61.82%	322.56	62.88%
	Total	584	100%	513	100.00%
Q12	**During the search phase, the business looks for a physical location for the business. Would you like the City of Gainesville to provide services during the search phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	249	43.08%	221.63	43.58%
	No	329	56.92%	286.89	56.42%
	Total	578	100%	508.52	100.00%
Q13	**During the shape phase, the business undertakes activities such as drawing up development/building plans. Would you like the City of Gainesville to provide services during the shape phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	235	41.30%	208.07	41.57%
	No	334	58.70%	292.53	58.43%
	Total	569	100%	500.6	100.00%

Q14	During the build phase, the physical location is built. Would you like the City of Gainesville to provide services during the build phase				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	253	45.18%	224.07	45.76%
	No	307	54.82%	265.61	54.24%
	Total	560	100%	489.68	100.00%
Q15	**The hire phase involves activities such as advertising for employees and hiring them. Would you like the City of Gainesville to provide services during the hire phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	174	31.02%	156.93	31.98%
	No	387	68.98%	333.82	68.02%
	Total	561	100%	490.75	100.00%
Q16	**During the taxes phase, the business puts bookkeeping procedures in place, hires an accountant, etc. Would you like the City of Gainesville to provide services during the taxes phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	180	32.20%	158.82	32.45%
	No	379	67.80%	330.68	67.55%
	Total	559	100%	489.5	100.00%
Q17	**The celebrate phase gives the business the opportunity to provide accolades to everyone that helped start the business. Would you like the City of Gainesville to provide services during the celebrate phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	226	40.72%	196.52	40.43%
	No	329	59.28%	289.66	59.57%
	Total	555	100%	486.18	100.00%
Q18	**During the open phase, the business has a grand opening, starts selling goods or services, etc. Would you like the City of Gainesville to provide services during the open phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	233	42.21%	206.43	42.62%
	No	319	57.79%	277.91	57.38%
	Total	552	100%	484.34	100.00%
Q19	**The grow phase starts the process of brainstorming again and moving forward with increasing the size of the business. Would you like the City of Gainesville to provide services during the grow phase?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	220	40.67%	190.31	39.82%
	No	321	59.33%	287.6	60.18%
	Total	541	100%	477.91	100.00%
	The next questions are about business interactions with the City				
Q20	**Often times, when establishing a business location, a search is done. In the last 12 months, did you conduct a search for this location?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	73	13.64%	66.9	14.23%
	No	462	86.36%	403.22	85.77%
	Total	535	100%	470.12	100.00%

Q21	**While you were searching for this location, what City of Gainesville departments did you interact with about this search?** *(Only asked to respondents who answered "yes" to previous question) **				
	Answer	**N**	**%**	**W_N**	**W%**
	Department of Doing (Planning and Building Inspection) - assists with understanding land management codes, issues building permits, etc.	18	25.71%	15.58	23.95%
	Public Works - issues permits (such as for stormwater), fixes public infrastructure, etc.	7	10.00%	6.88	10.58%
	Gainesville Regional Utilities - turns on electricity, offers internet service, trims trees away from power lines, etc.	24	34.29%	20.88	32.10%
	Budget and Finance - issues business licenses	16	22.86%	14.09	21.67%
	Community Redevelopment Agency - provides redevelopment assistance and incentives	4	5.71%	3.39	5.21%
	Other (please specify)	4	5.71%	2.88	4.43%
	None	33	47.14%	33.91	52.15%
	Total	70		65.04	

Q22	**In the last 12 months, did you draw up development/building plans for this location?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	34	6.38%	34.15	7.29%
	No	499	93.62%	434.19	92.71%
	Total	533	100%	468.34	100.00%

Q23	**While you were drawing up development/building plans for this location, what City of Gainesville departments did you interact with about the development/building plans?** *(Only asked to respondents who answered "yes" to previous question) **				
	Answer	**N**	**%**	**W_N**	**W%**
	Department of Doing (Planning and Building Inspection) - assists with understanding land management codes, issues building permits, etc.	22	68.75%	21.53	67.24%
	Public Works - issues permits (such as for stormwater), fixes public infrastructure, etc.	13	40.63%	13.16	41.11%
	Gainesville Regional Utilities - turns on electricity, offers internet service, trims trees away from power lines, etc.	16	50.00%	14.45	45.11%
	Budget and Finance - issues business licenses	6	18.75%	4.63	14.45%
	Community Redevelopment Agency - provides redevelopment assistance and incentives	4	12.50%	2.13	6.64%
	Other (please specify)	1	3.13%	1.27	3.97%
	None	6	18.75%	6.95	21.71%
	Total	32		32.03	

Q24	**In the last 12 months, did you build this location?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	23	4.34%	24.26	5.21%
	No	507	95.66%	440.92	94.79%
	Total	530	100%	465.18	100.00%

Q25	**While you were building this location, what City of Gainesville departments did you interact with about building this location?** *(Only asked to respondents who answered "yes" to previous question)* *				
	Answer	**N**	**%**	**W_N**	**W%**
	Department of Doing (Planning and Building Inspection) - assists with understanding land management codes, issues building permits, etc.	16	69.57%	17	70.10%
	Public Works - issues permits (such as for storm water), fixes public infrastructure, etc.	12	52.17%	14.75	60.81%
	Gainesville Regional Utilities - turns on electricity, offers internet service, trims trees away from power lines, etc.	16	69.57%	16.78	69.19%
	Budget and Finance - issues business licenses	7	30.43%	8.23	33.93%
	Community Redevelopment Agency - provides redevelopment assistance and incentives	4	17.39%	2.19	9.03%
	Other (please specify)	0	0.00%	0	0.00%
	None	4	17.39%	4.3	17.73%
	Total	23		24.26	

Q26	**On a scale of 0 to 10 where 0 is the worst possible service, and 10 is the best possible service, please rate the quality of service you received from the Department of Doing (Planning and Building Construction).** *(Only asked to respondents who answered "Department of Doing" for Q21, Q23, or Q25)* *				
	Answer	**N**	**%**	**W_N**	**W%**
	0	0	0.00%	0	0.00%
	1	2	5.00%	2.09	5.38%
	2	3	7.50%	3.77	9.71%
	3	4	10.00%	4.15	10.70%
	4	1	2.50%	2	5.15%
	5	2	5.00%	1.17	3.03%
	6	2	5.00%	3.48	8.96%
	7	6	15.00%	4.4	11.33%
	8	5	12.50%	3.81	9.81%
	9	5	12.50%	5.42	13.96%
	10	10	25.00%	8.53	21.97%
	Total	40	100%	38.82	100.00%

Q27	**On a scale of 0 to 10 where 0 is the worst possible service, and 10 is the best possible service, please rate the quality of service you received from the Public Works department.** *(Only asked to respondents who answered "Public Works Department" for Q21, Q23, Q25)* *				
	Answer	**N**	**%**	**W_N**	**W%**
	0	0	0.00%	0	0.00%
	1	1	5.00%	1.32	5.75%
	2	0	0.00%	0	0.00%
	3	1	4.76%	1.06	4.62%
	4	1	4.76%	1.05	4.57%
	5	4	19.05%	6.4	27.86%
	6	2	9.52%	1.4	6.10%
	7	2	9.52%	1.67	7.27%
	8	1	4.76%	1.06	4.61%
	9	2	9.52%	2.22	9.66%
	10	7	33.33%	6.79	29.56%
	Total	21	100%	22.97	100.00%

Q28	**On a scale of 0 to 10 where 0 is the worst possible service, and 10 is the best possible service, please rate the quality of service you received from Gainesville Regional Utilities.** *(Only asked to respondents who answered "Gainesville Regional Utilities" for Q21, Q23, Q25) **				
	Answer	**N**	**%**	**W_N**	**W%**
	0	0	0.00%	0	0.00%
	1	4	10.26%	5.02	13.68%
	2	1	2.56%	0.29	0.79%
	3	0	0.00%	0	0.00%
	4	1	2.56%	0.91	2.48%
	5	8	20.51%	7.04	19.18%
	6	4	10.26%	3.44	9.38%
	7	7	17.95%	7.95	21.67%
	8	5	12.82%	4.07	11.09%
	9	5	12.82%	4.51	12.29%
	10	4	10.26%	3.46	9.44%
	Total	39	100%	36.69	100.00%

Q29	**On a scale of 0 to 10 where 0 is the worst possible service, and 10 is the best possible service, please rate the quality of service you received from the Budget and Finance department.** *(Only asked to respondents who answered "Budget and Finance Department" for Q21, Q23, Q25) **				
	Answer	**N**	**%**	**W_N**	**W%**
	0	1	4.35%	0.26	1.25%
	1	0	0.00%	0	0.00%
	2	1	4.35%	1.05	5.07%
	3	1	4.35%	1.02	4.93%
	4	0	0.00%	0	0.00%
	5	4	17.39%	3.22	15.54%
	6	0	0.00%	0	0.00%
	7	4	17.39%	4.94	23.84%
	8	3	13.04%	1.94	9.36%
	9	6	26.09%	5.11	24.66%
	10	3	13.04%	3.18	15.35%
	Total	23	100%	20.72	100.00%

Q30	**On a scale of 0 to 10 where 0 is the worst possible service, and 10 is the best possible service, please rate the quality of service you received from the Community Redevelopment Agency.** *(Only asked to respondents who answered "Community Redevelopment Agency" for Q21, Q23, Q25) **				
	Answer	**N**	**%**	**W_N**	**W%**
	0	0	0.00%	0	0.00%
	1	0	0.00%	0	0.00%
	2	0	0.00%	0	0.00%
	3	0	0.00%	0	0.00%
	4	1	14.29%	0.33	7.55%
	5	0	0.00%	0	0.00%
	6	0	0.00%	0	0.00%
	7	1	14.29%	1.27	29.06%
	8	2	28.57%	0.76	17.40%
	9	1	14.29%	0.44	10.07%

	10	2	28.57%	1.57	35.92%
	Total	7	100%	4.37	100.00%
Q31	**During which of these 3 phases would you like the City of Gainesville to improve services? ***				
	Answer	**N**	**%**	**W_N**	**W%**
	Search - business looks for a location	20	31.74%	15.05	26.10%
	Shape - business draws up development/building plans	17	26.99%	15.4	26.71%
	Build - business builds the physical location	19	30.16%	16.72	28.97%
	None of the above	26	41.27%	25.61	44.40%
	Total	63		57.18	
Q32	**Were there services you were seeking from the City of Gainesville, but unable to find, during any of these 3 phases? Please select all that apply. ***				
	Answer	**N**	**%**	**W_N**	**W%**
	Search - business looks for a location	5	8.62%	4.58	8.08%
	Shape - business draws up development/building plans	9	15.52%	9.4	16.58%
	Build - business builds the physical location	4	6.90%	4.03	7.11%
	None of the above	44	75.86%	38.67	68.23%
	Total	62		56.68	
Q33	**How likely are you to recommend opening a business in Gainesville to others?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Extremely likely	72	13.66%	64.29	13.89%
	Very likely	159	30.17%	136.96	29.59%
	Moderately likely	173	32.83%	150.6	32.53%
	Slightly likely	61	11.57%	52.76	11.40%
	Not at all likely	62	11.76%	58.29	12.59%
	Total	527	100.00%	462.9	100.00%
Q34	**You said you are [Extremely, Very, Moderately, Slightly, not at all likely] to recommend opening a business in Gainesville to others. Did your experience with the City of Gainesville influence your answer?** *(Only asked if Q33 was answered)*				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	245	47.02%	213.68	46.55%
	No	276	52.98%	245.33	53.45%
	Total	521	100.00%	459.01	100.00%
Q35	**How would you rate the success of your business at this location?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Very Successful	232	44.62%	209.62	45.82%
	Somewhat Successful	253	48.65%	211.47	46.22%
	Somewhat Unsuccessful	27	5.19%	29.46	6.44%
	Very Unsuccessful	8	1.54%	6.93	1.52%
	Total	520	100.00%	457.48	100.00%

	You said your business has been Very/Somewhat Successful/Unsuccessful. What are the reasons? * *(Only asked if Q35 was answered)*				
	Answer	**N**	**%**	**W_N**	**W%**
	The actions of the business owners and employees	311	61.46%	277.77	62.23%
	The state of the local economy	201	39.72%	177.47	39.87%
	Prior experience with a similar business	137	27.07%	120.11	26.98%
	Interactions with the City of Gainesville	39	7.71%	37.24	8.37%
	Other (please specify)	117	23.12%	97.91	21.99%
	Total	506		445.148	
Q37	**In the last 5 years, have you done any remodeling at this location?**				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	145	27.99%	140.8	30.98%
	No	373	72.01%	313.71	69.02%
	Total	518	100.00%	454.51	100.00%
Q38	**Was this remodeling done for the purpose of business expansion?** *(Only asked if the previous question was answered "yes")*				
	Answer	**N**	**%**	**W_N**	**W%**
	Yes	60	41.38%	61.63	43.77%
	No	85	58.62%	79.16	56.23%
	Total	145	100.00%	140.79	100.00%
	The last six questions of the survey collected quantitative financial data which have been omitted to preserve the privacy of the respondents and could not be aggregated meaningfully.				
Q39	**For the entire year of 2016, approximately how much were the revenues for this location? Please give your answer in dollars. Your best guess is fine.**				
Q40	**For the entire year of 2016, approximately how much were the expenses for this location? Please give your answer in dollars. Your best guess is fine.**				
Q41	**The capital investment is the money spent by a company to acquire or upgrade physical assets such as property, industrial buildings or equipment. It is often used to undertake new projects or investments.** **For the entire year of 2016, approximately how much was spent on capital investment for this location? Please give your answer in dollars. Your best guess is fine.**				
Q42	**For the first half of 2017, approximately how much were the total revenues for this location? Please give your answer in dollars. Your best guess is fine.**				
Q43	**For the first half of 2017, approximately how much were the total expenses for this location? Please give your answer in dollars. Your best guess is fine.**				
Q44	**The capital investment is the money spent by a company to acquire or upgrade physical assets such as property, industrial buildings or equipment. It is often used to undertake new projects or investments.** **For the first half of 2017, approximately how much was spent on capital investment for this location? Please give your answer in dollars. Your best guess is fine.**				