INVESTING IN HOTEL ANALYTICS

The factors to consider as part of an hotelier's investment in distribution analytics

March 2019

CONTRIBUTORS

CO-CHAIRS AND AUTHORS

Josh Galun	Hospitality & Travel Lead at Excella Consulting
Nitin Bajaj	Director. Business Intelligence and Analytics at NTTData
Matthew Goulden	CEO at Triometric

GROUP CO-ORDINATION

Sonja Woodman	Marketing Manager at Triometric

CONTRIBUTING AUTHORS

Audrey Murante	Director, Global Alliances & Initiatives at IDeaS - A SAS Company
Nicholas Tsabourakis,	Director at Bespoke Revenue Management
Shane Forster	Head of Marketing at Customer Alliance
Casey Davy	Vice President Sales and Head of Marketing at Busy Rooms
Lori Grant	Product Delivery Manager at IHG
Connie Fuchs	Senior Manager Business Analysis & Performance at Worldhotels
Dave Chestler	HEDNA Board Member

EDITING ASSISTANCE

Rajesh Vohra	Director at Sarova Hotels

HEDNA Board Liaison

Lew Harasymiw	Director, Distribution Strategy - Connectivity at IHG
Jodie Gibson	Manager Strategic Accounts at DerbySoft

DEFINITIONS

The following is a list of definitions of common terms used throughout this paper.

API
: Application Programming Interface is a mechanism to use programming code to execute functions in a piece of software.

Aggregator
: A data provider that supplies availability and rates from multiple hotel suppliers. Supplier is the Universal API-standard term for a hotel vendor such as a hotel chain, hotel property, or hotel brand name.

Channel Manager
: A system that allows the hotelier to quickly and efficiently manage room rates, availability, and restrictions across all their connected online distribution channels.

Data Access
: The act or method of viewing or retrieving stored data

Database
: A digital collection of data and the structure around which the data is organized.

JSON
: JavaScript Object Notation is a lightweight data-interchange format. It is easy for humans to read and write. It is easy for machines to parse and generate.

Data Integration
: The process of combining data from different sources and presenting it in a single view.

Data Quality
: The measure of data to determine its worthiness for decision making, planning, or operations

Data Source
: Any provider of data–for example, a database or a data stream.

GDS
: Global Distribution System – a computerized reservation networks through which users (travel agents, airline employees, travelers) view data on a wide range of travel services, including air travel, hotel, car rental and like services.

KPI
: Key Performance Indicator

OTA
: Online Travel Agent

PMS
: Property Management System

PII
: Personally identifiable information is any data that could potentially identify a specific individual.

PCI-DSS
: Payment Card Industry Data Security Standard is an information security standard for organizations that handle branded credit cards from the major card schemes.

RMS
: Revenue Management System

ROI
: Return on Investment

XML
: Extensible Mark-up Language (XML) is a universal data format used for representation and transfer of structured data on the web or between different applications.

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY

Hotel companies are swimming in data. From the property management system to the hotel website, there are a plethora of systems capturing data that spans the gamut from online customer behavior to sales and marketing costs. On top of this, third-party data providers give hoteliers access to market and customer data that was never-before available.

This explosion of data has set off a revolution in distribution analytics. Hotel companies can achieve new levels of precision in their channel mix to optimize occupancy, rate, and profits. In today's marketplace, distribution analytics are the hotelier's competitive edge.

But while hotel companies worldwide see the value in distribution analytics, many are unsure how to adopt it. Distribution analytics requires an investment in technology, including new tools and platforms that hotel companies may not yet have experience in. Similarly, distribution analytics requires staff who have the skills to derive real value from these tools, and those skills are not yet common within the industry.

Every hotel company is different, so there is no one-size-fits-all solution. This white paper explains key areas that hotel companies should evaluate before making an investment in distribution analytics. In so doing, hoteliers will better be able to define a vision for distribution analytics at their company and plan the necessary investment needed to accomplish it.

2. INTRODUCTION

The HEDNA Analytics Working Group was founded at the HEDNA European conference in Dublin in June 2017 in response to member interest for information and practical guidance to using data within their organizations. The first phase of the Group's work was to survey over 1,000 hoteliers on the current landscape of data analytics in the industry. The Survey Findings were presented at the HEDNA Global Distribution Conference in Austin in January 2018 and published in March 2018.

This laid the foundation for the second phase of the Working Group's activity, which is to document the considerations for making investment decisions in analytics capabilities. This white paper provides an overview of the key issues and requirements for adopting analytics to optimize a hotel company's distribution operations.

The final phase of the Working Group's mission will be to augment these important considerations with best practices and recommendations. The final phase and this document will be presented at HEDNA's next European conference in Madrid in May 2019.

3. OVERVIEW

In recent years the hotel industry has experienced dramatic changes, particularly in the area of hotel channel distribution. In order for hotels to navigate this shifting distribution paradigm, and optimize their channel mix, they must embrace analytics to give them clearer visibility into channel costs, short-term effectiveness (at moving inventory), long-term effectiveness (at creating loyal guests), and Return on Investment (RoI).

Emerging technologies are increasingly changing the practice of hotel distribution. Operational reports and comp sets are being supplemented with advanced analytical capabilities and techniques. The art of distribution is increasingly giving way to the science of distribution.

There is tremendous opportunity. Hoteliers can use previously unavailable data to achieve new heights of channel optimization. They can improve their insight into channel costs, to improve profitability along with the top line revenue. They can establish clear, data-driven guidance that improves staff decision-making and reduces the impact of turnover.

This also means that hoteliers must embrace the data-driven age and adopt analytics just to stay competitive. Those hoteliers who continue to work from Excel spreadsheets will be at a disadvantage to their competitors.

Implementing analytics, especially beyond departments and across the organization can be a daunting task. Because it is a new area to many hoteliers, it can be hard to know where to begin, how to go about planning it, what approach is best, and how much needs to be invested. In addition, analytics can require technical and mathematical capabilities not common within the industry.

There is far too much variety amongst hotel companies for there to be a one-size-fits-all answer. However, there is a planning process and a common set of considerations that can help hoteliers determine the right approach for them. This white paper describes a set of relevant factors to take into account as they consider options for investing in improving their distribution analytic capabilities.

4   DEFINING DISTRIBUTION ANALYTICS

To begin, it helps to map out the key objectives, and then unify employees around a shared definition of what distribution analytics means to the organization. An analytics project cannot be driven by a single person; to achieve success it requires commitment from key stakeholders. This includes input from the departmental users to define their requirements, and buy-in from the staff who are required to use the systems being implemented. Therefore, it is important that everyone agrees on the meaning of the core concept of analytics and how it can benefit their hotel operations. The term "analytics" is so frequently and commonly used that this might seem unnecessary. But in practice, it remains a fuzzy concept for many – especially those whose day-to-day jobs have not yet been impacted by analytics to date. Ask your co-workers what they think it means, and you're likely to hear a very wide range of definitions.

Formally, analytics is the practice of the discovery, interpretation, and communication of meaningful patterns in data to support decision-making. This means that analytics is not just an enhanced form of operational or financial reporting. Reporting only becomes analytics when it supports the discovery of meaningful patterns in support of decision-making:

Pattern discovery is often divided into four types:

- Descriptive Analytics - answers the question of what happened?
- Diagnostic Analytics - answers the question of why did X happen?
- Predictive Analytics - answers the questions of what do we think will happen?
- Prescriptive Analytics - answers the questions of what should we do to eliminate a future problem or take advantage of a positive trend?
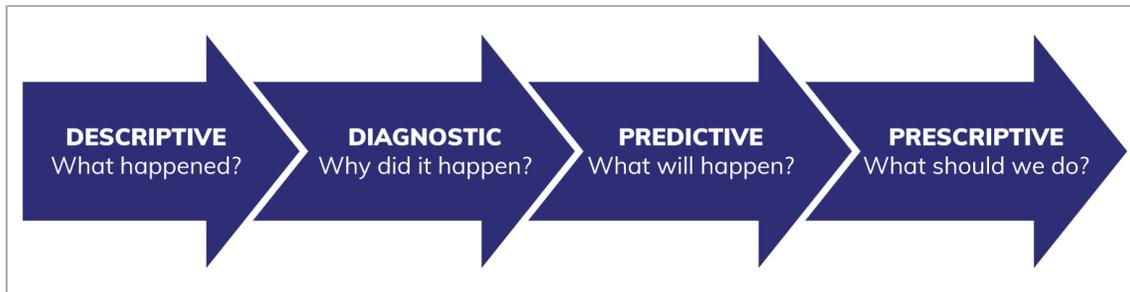
Figure 1: The 4 Types of Analytics

With the focus on pattern discovery, the "interpretation" and "communication" elements of analytics should not be neglected. After all, new findings are pointless if they are not in the hands of staff, understandable, and actionable. In fact, many hotel companies already have access to large amounts of data, and could get huge benefits simply from making the data easier to access and analyze, better visualized, more widely distributed, and more interactive.

This document specifically deals with "distribution analytics." This means applying analytics techniques for the specific purpose of planning the most effective and profitable distribution strategy for hotel inventory. It involves optimizing the channel mix based on cost of acquisition, rates, market demand, availability, competition, and customer segmentation.

In the original HEDNA Hotel analytics working group charter some key areas of distribution analytics were identified:

- **Distribution Channel Performance** – Hotels need to appreciate the relative costs of direct and indirect distribution channels.
- **Market Trends** – In a climate of high look to book ratios, hotels can lack insight into the spectrum of which products / destinations are being searched for, by channel and how they convert.
- **Distribution Supply Insight** – Hotels need to appreciate the relative performance of the different channels of the segments they serve, for example GDS vs OTAs, business vs leisure.
- **Operational Performance** – the on-line travel market revolves around speed and accuracy. Visibility is needed in real-time IT operations and connectivity performance to ensure opportunities are not missed.

As can be seen from the list above, a distribution strategy is one of the more complex challenges for an hotelier, because it hinges upon so many variable items. Many different kinds of data and methods of analysis are relevant to this process. Specific examples of distribution analytics include the following:

- Predicting demand for a specific competitor within an hotelier's market segment
- Analyzing search traffic for specific products coming from the various channels
- Optimizing pricing to balance rate and occupancy in order to maximize an hotelier's RevPAR
- Determining the exact cost of acquisition for a specific channel
- Assessing whether there is a difference in ancillary spend per channel
- Ascertaining the lifetime value of a new customer based on acquisition channel
- Analyzing the interplay between group sales and leisure shoulder dates
- Examining the relationship between hotel website activity and demand on other channels

There are many factors that impact distribution decision-making, the scope of distribution analytics is wide. Nevertheless, there are forms of hotel analytics that fall outside of scope. For example:

- The analysis of guest satisfaction surveys to determine capital projects for hotel renovations.
- Studying the performance of a hotel's restaurant or room service.
- Measuring the usage of the hotel's mobile key app.

It can be helpful to give staff a clear sense of these boundaries. Other forms of hotel analytics may require different data sets and different systems, and for many hoteliers it would be overly ambitious to tackle all forms of hotel analytics concurrently.

4. ESTABLISHING AN ANALYTICS VISION

Using this definition of distribution analytics as a foundation, hoteliers should work with their departmental heads (such as Revenue Managers, Channel Managers, etc.) to set a common vision for how they will use distribution analytics to support both their strategy and their day-to-day business processes. Since distribution analytics is a new discipline for many hoteliers, this will provide the opportunity to establish the essential framework for implementing the agreed systems and processes as part of any implementation.

A good way to establish this vision is by developing a set of simple use cases. These will have the information that hoteliers need to estimate investments and plan the project. They will establish the parameters needed to assess software tools, select technical infrastructure, and build staff skillsets. Finally, they will provide a backlog of prioritized analytics capabilities to build when the project is launched.

To be clear, the vision and the use cases should not be set in stone. In fact, hoteliers can expect the use cases to be refined, updated, and changed as employees learn through practice. Some use cases will turn out to be too challenging or to have less value than expected. On the other hand, fruitful new ideas will arise as staff grow comfortable with analytics and see new possibilities.

For this reason, we do not recommend a "big bang" approach to analytics implementation. It would be ambitious for hoteliers to assume they have that perfect understanding of their distribution analytics needs from day one. Hoteliers should plan for a journey towards greater analytics maturity, not a single leap but in series of well-planned steps.

Instead, we recommend an iterative and incremental process, adopting new analytics in small pieces. This will allow your staff to work with new analytics capabilities early and often, and provide feedback that keeps the project delivering the maximum possible value and impact.
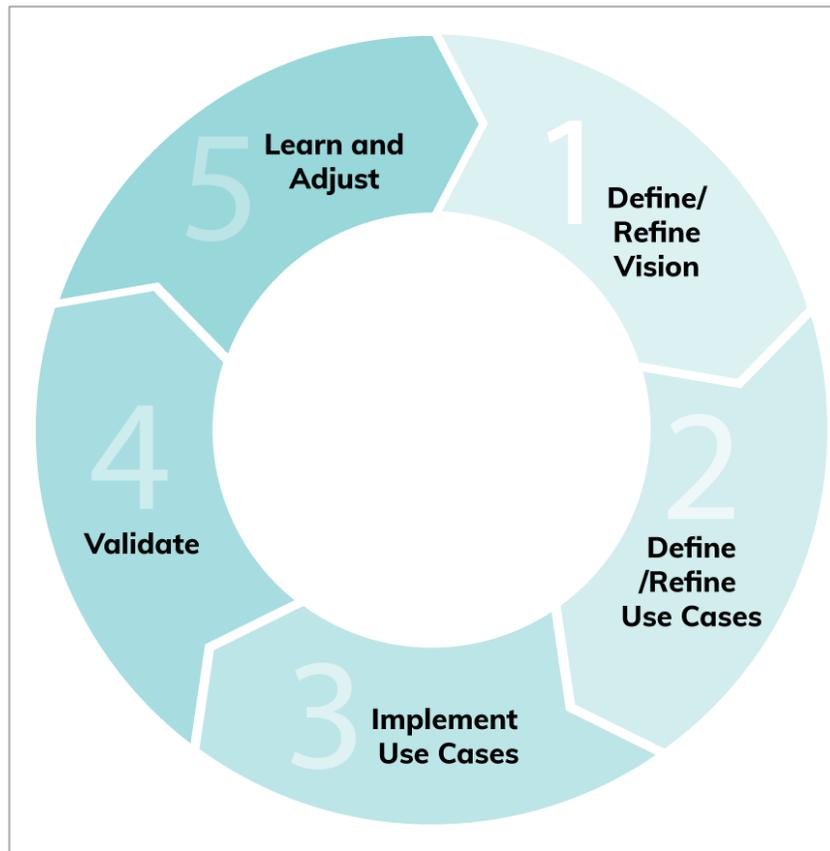
Figure 2: Iterative Analytics Adoption Steps

The combination of initial use cases with an incremental approach strikes an appropriate balance between planning and flexibility. There will be sufficient clarity to manage the project and measure return on investment, while still giving space for learning and adjustments in a rapidly-changing area of hospitality.

5. GATHERING USE CASES

The use cases should capture the business question being answered, the general analytics approach that would be needed, and the data sets that would underpin the analytics.

A good way to establish use cases is by starting with the business process and then drilling into them to find opportunities for analytics. In particular:

- Identify the processes that are being executed within your business on a regular basis.
- Determine the decisions that are made within that business process.
- Define the key performance indicators (KPIs) that help you make those decisions.
- Determine which of the four pattern discovery approaches (Descriptive, Diagnostic, Predictive, and Prescriptive) would provide KPIs to support those decisions. A single KPI may have multiple analytics approaches.
- Describe the analytics application at a high level. E.g., "a report that shows our occupancy rate for the past 30 days" or "a model that suggests room pricing based on our comp set."
- Assess how the analytics would need to be interpreted and communicated in order to be applied by staff.
- Assign a ranked prioritization to the use cases. This could be based on estimated ROI or on more subjective measures. This will help determine which use cases are worth implementing within project budget and which will be implemented first.

1	Business Processes	5	Analytics Approach
2	Major Decisions	6	Communication Requirements
3	Key Performance Indicators	7	Prioritize
4	Pattern Discovery Approach		

Figure 3: Factors in Developing Analytics Use Cases

Hoteliers may find it challenging to identify use cases in an area that is brand new to them. Therefore, we recommend a multi-pronged approach to defining them. Here are six distinct steps as part of the critical planning and preparation process:

Figure 4: 6 Steps for Preparing Use Cases

1.   **Conduct Interviews** with a variety of team members and departments - ranging from executives to managers to analysts. Spend time with more than one stakeholder to gain insight on what their reporting and data needs are:

    i.   Begin by inquiring about their high-level business objectives. Ask what their success metrics are, how these are monitored and reported, and what types of challenges are preventing them from meeting their business objectives.

    ii.   Ask about things they would change, items they would like, and "what if" scenarios, to name a few.

    iii.   Learn what types of routine analysis your stakeholders currently perform, what types of data is currently being used, how this data is obtained, and what's done with the information once they get it.

2.   **Observe or shadow internal user behavior** in your current reporting environment. This is a useful technique to detect any bottlenecks within the current state. It's often hard for a user to communicate obstacles, so this can also be helpful in discovering overlooked requirements that may have initially been missed.

3. **Inventory reports** to determine if there are opportunities for improvement. (In addition, if you implement a new centralized analytics tool, it is likely that you may recreate existing reports in the new tool.) Gather a list of the reports you are currently using and what purpose they serve and determine what personnel and skill sets are supporting each.

4. **Inventory current systems** – such as your PMS, RMS, and channel manager – to identify reports that you are not currently using. Often, hoteliers get used to working with their standard set of reports, and haven't explored the additional capabilities that are already available to them. There may be opportunities to use these tools in greater ways, thereby avoiding the need to implement new tools. These unused capabilities may also provide helpful starting points to identify business processes, KPIs, and analytics approaches that you had not yet considered.

5. **Research industry analytics products** to identify what reports and analytics capabilities they offer. Again, their capabilities may provide useful ideas to identify business processes, KPIs, and analytics approaches that you had not yet considered. In addition, this can provide a head start on identifying tools that you may want to use as part of your implementation project.

6. **Attend industry events** and tradeshows to speak to peers from outside of your own company. This can add valuable perspective and fresh ideas to your own internal use case analysis. They can assist hoteliers in uncovering reporting and data capabilities they may not even know exist. Staying abreast of the technology trends as well as new innovations and possibilities on the horizon within the industry will ensure that hoteliers are ahead of the curve and do not fall behind their competitors.

6. DATA INTEGRATION

As you make progress in gathering use cases in your organization, it is important to start creating the data layer on which the use cases will be built. Outside of very simple descriptive analytics – such as trend analysis of booking volumes or revenue reporting – almost all use cases will require integrating data from multiple source systems.

There are five primary factors to be considered to assess the level of data integration that needed:

- **Data Sources:** As the number of source systems increases, the complexity levels of the data integration will increase, requiring tradeoffs in either the frequency or in the volumes of data that can be processed. Additionally, heterogeneity in the format in which source systems store data increases the build complexity of the data integration. For example, date formats are handled differently in Excel, relational databases, file extracts, and in big data technologies. Wrangling this source data into a common format requires building and testing transformations for each source system, increasing the time to value of that data for the person using that data

- **Analytical Relationships:** To what extent do individual use cases require analysis or visualization from varied data across multiple data sources? The more fully that data needs to be tied together across data sources, the greater the effort needed to achieve data integration between the systems.

- **Data Frequency:** How frequently does the use case need to be executed to support the hotelier's business processes, and therefore, how often does it need data from the source systems? For many distribution use cases, the data is only needed on a periodic basis (monthly, weekly, daily, etc.). But some use cases require real-time data. As frequency increases and particularly once the data integration becomes real-time, integrations tend to become more complex, more brittle, and less reliable over time. The greater the frequency, the more time and effort will need to be invested to maintain the performance and quality of the data integration.

- **Data Volume:** Another important factor to consider is the volume of data in the source systems that needs to be consolidated. The good news is that storage has become incredibly cheap, and few hoteliers will have anywhere near the volume that storage space becomes an issue. But volume does still have an impact. The higher the volumes, the greater the load on the source systems to service the requests. Hoteliers may need to invest in more computing power, or accept a lower data integration frequency. Similarly, if the analytics use cases require scanning high volumes of data, this increases the load during use case execution. Again, this may require hoteliers to invest in more computing power or run analytics less frequently.

- **Data Identifiers and Granularity:** Each source system may have its own scheme for identifying the data and may store data at different levels of granularity (detail). For example, the booking system may identify a room reservation as an alphanumeric code and store it as a single record, but the financial system may identify the same reservation as a numeric identifier, and store one record per GL-code. As data is integrated across the system, a common identifier needs to be identified and chosen, and the final granularity of the data determined based on the use case.

Data integration is a mature technical discipline and there is a plethora of tools and technologies available to develop data integrations. For an hotelier evaluating their options for data integration platforms and tools, the following considerations are important for making selecting the right tool:

- Licensing schemes: Broadly there are three categories to consider:
    - i) Perpetually licensed products with an annual maintenance fee,
    - ii) Software as a Service (SaaS) products with a bundled monthly license management fee
    - iii) Open source products that are free to download and use.

- Metadata management capabilities can help hoteliers keep the data clean and well understood by mapping data sources to use cases, clearly visualizing which data sources support which data integrations, documenting data definitions, and tracking the types of transformations that data undergo as they are integrated.

- Ability to handle big data platforms has become an important consideration for data integration tools, as the volume, velocity and variety of data in source systems increases exponentially.

- Most data integration tools provide out of the box connectors for common data sources such as relational databases, Excel files, CSV files, etc. Data integration tools may also provide the ability to connect to additional data sources such as Hadoop, SalesForce.com, mainframe systems, and other specialized applications. Hoteliers should evaluate which data connectors they need.

- Market share, service firm partnerships, and availability of skilled resources, are factors that should be weighted in the selection process, to ensure that you can successfully hire or contract the resources you need to build out your platform.

7. DATA QUALITY

Analytics requires high quality data to deliver meaningful insights. If the source data is unreliable, inconsistent, or incomplete then the insights delivered will be of minimal use. Hoteliers often struggle to have on-property staff maintain accurate and complete data. These staff have high turnover rates and may not have any connection to why the data quality is important. These risks can be mitigated by:

- Training staff on the relevant and various tools available to them. A well-trained staff is more confident and more motivated to ensure all data entry and logging is done correctly.
- Educating staff in the value, impact, and importance of data quality to the hotel and the company. Often, staff simply do not understand why it matters.
- Providing staff with incentives to achieve and maintain higher levels of data quality.
- Developing and implementing clear and simplified handover processes amongst hotel staff for operational activities and need for logging handovers in relevant systems
- Developing system automations/templates that speed up data entry and logging while ensuring that data is consistent across multiple staff roles and across multiple properties.

Evaluating the state of data quality is an important first step in building the foundation for analytics within your enterprise. As you evaluate data quality, ask yourself: Is the data complete? Does it have unique identifiers allowing it to match external file to main dataset? How high is the risk of mismatches between the unique identifiers? Is the data itself consistent? (E.g. are dates always formatted in the same format? Or are a number fields always filled with numbers and not accidentally with text?)

On the flip side, also consider the quality of data in relation to the business use case that you are solving for. For example, reporting gross revenue on a quarterly basis to meet regulatory and compliance requirements requires a very high degree of data quality, for each accounting GL-code for each reservation in that timeframe. However, if you are trying to determine conversion rate by channel for your online reservations, some leeway at the atomic data level may be tolerable, while still yielding actionable insights for the marketing team.

We recommend establishing a formal framework for managing data quality across your analytics use cases, to track the entire data flow and proactively manage data quality on a continuous basis. Otherwise, your data quality may gradually diminish over time.

The following diagram presents key steps and activities required to establish such a framework.
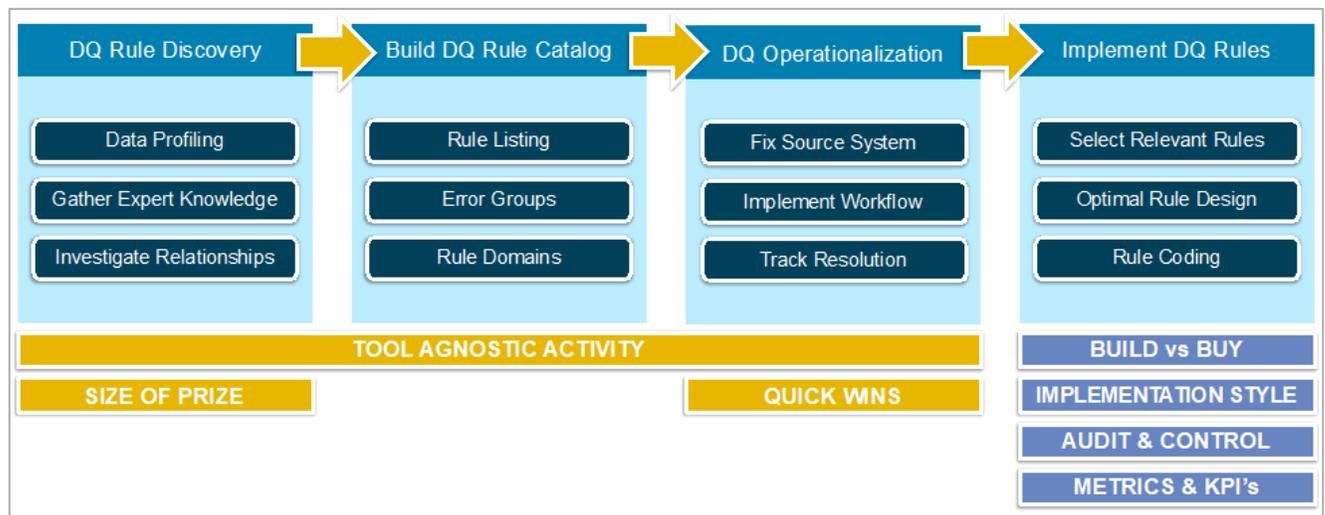

DQ Rule Discovery	Build DQ Rule Catalog	DQ Operationalization	Implement DQ Rules
Data Profiling	Rule Listing	Fix Source System	Select Relevant Rules
Gather Expert Knowledge	Error Groups	Implement Workflow	Optimal Rule Design
Investigate Relationships	Rule Domains	Track Resolution	Rule Coding

TOOL AGNOSTIC ACTIVITY

SIZE OF PRIZE

QUICK WINS

BUILD vs BUY

IMPLEMENTATION STYLE

AUDIT & CONTROL

METRICS & KPI's

Figure 5: Data Quality Framework

The first three phases of this framework are agnostic of the tooling or platform on which the data quality framework will be implemented. In these first three steps, the primary participants are business analysts and data analysts in the organization, supported by business stakeholders that have expert knowledge of specific areas in the distribution value chains. As you perform data profiling and investigate relationships between data in your organization, you will be able to quantify the "size of prize" of implementing the data quality framework, an important gate check to ensure you are not over (or under) investing in this domain.

As you move towards operationalizing the data quality rules, the general rule of thumb is to prioritize areas of with high impact and minimal change complexity, and move down the impact and up the change complexity continuum, measuring ROI at each step and stopping when the effort expended on the change is greater than the realized benefit.

Finally, consider implementing a data quality tool to store and implement these rules, ensuring that your enterprise does not accumulate data debt over time that makes the value of your analytics insights suspect. As you consider tool selection for implementing, managing and monitoring data quality rules, consider factors that are described in the data integration section above.

8. DATA STORAGE

Data storage refers to storing and managing data through the entire information management lifecycle. It includes data that are created by operational systems, data that are derived based on calculations and transformations, and third-party data obtained from vendors or aggregators.

Estimating the volume of data that will be stored involves calculating the current size of the data that will be integrated, the rate of growth, and the duration for which data must be stored. You should also remember that your data integration may include new combined, hybrid, aggregated, or pre-calculated fields that do not exist in the source systems. This contributor to data volume is often forgotten during planning.

Also, you should include a contingency or "flex" amount to account for unknown factors or unexpected changes to source systems and use cases. The industry norm is to add an additional 30% to your initial calculation.

But data storage is more than just disk space. A robust framework for data storage will define, implement and monitor all aspects of data security, data privacy, and data retention in the enterprise.

Generally speaking, most distribution analytics use cases should not require personally identifiable information (PII) or payment card industry (PCI) data. For example, consider a use case to evaluate the impact of changing market conditions on each channel. You will need to look at the aggregate booking rate in each channel, but there is no need to know the name of each person who booked or their credit card number. In fact, if there is an analytics use case that initially seems to require this information, it is worth reevaluating it to determine if there is a way to achieve the same analysis without that data.

Rather than automatically loading and storing all data from the source systems, it should be selected on a "need to know basis." If sensitive data is not required for the analysis/analytics system, then strip it our early in the process.

A good example of this in practice is Triometric's distribution analytics systems which capture and analyze API (XML/JSON) data. The primary objective of the product is to optimize the hotel's conversions through its third-party channels. However, the API data can include PII (the consumer) and PCI (their card details) which is not required for this purpose, so to nullify any risk this data is deleted at capture by the Triometric system.

The data selected to load into your analytics data should be linked to a specific use case. That way, you know exactly what purpose the data serves and what benefits it will create before you choose to bear the cost of storage, the impact to data speed, and the potential security/privacy risks.

Data security and privacy requirements can't be eliminated altogether. Even if you limit the data that you are storing to only what's necessary, no business can afford to let its data be exposed to its competitors and to the public. Therefore, before investing in distribution analytics, hotel companies should consider where they would store the data and how they would keep that data secured. This doesn't only mean where the data is stored, but who will be maintaining the environment to make sure that it stays up-to-date with security patches and security best practices.

Many hotel companies do have their own computing environments where an analytics database could be housed. But few have an IT operations staff with expertise in environment maintenance. Hotel companies should evaluate whether they are prepared to hire the necessary staff or contract with an outside vendor for environment support. Alternatively, the hotel companies may want to use cloud-based analytics software that stores the data in the vendor's environment, so that the hotel company doesn't need to take on that responsibility itself. In that case, the vendor's security processes and audit results should be validated before committing to the vendor's environment.

9. DATA ACCESS

When implementing an analytics platform, it is critical to consider how users will access the underlying data. Specifically, you must answer the questions:

- Who can access the data?
- What data can they access?
- Can users access the underlying data or do they use pre-built models?
- What tools can they use to access the data?
- Where can users work with the data?

The intuitively appealing approach is to grant full, direct database access to everyone in the company, and let them use the data wherever and however they choose. After all, you have gone to the trouble of integrating data, and it only becomes useful when people perform analytics with it.

But just like data storage, there are regulatory, security, and privacy considerations to consider for data access as well. In addition, there is also a trade-off between providing access and making sure that the data is being used and interpreted correctly. If not taken into account and mitigated, there is a potential for inexpert users to create queries which result in overly large processor overhead when they have full access to the underlying data.

What Data Can Users Access?

As discussed in the Data Storage section, some data may be subject to PII or PCI regulations. This data should be minimized or eliminated to the greatest extent possible.

However, not all data is as simple to categorize as PII and PCI. For example, what about commercially sensitive data such as the value of the total bookings through a distribution partner? Clearly this is important information to the commercial teams, but are we happy for other teams or users to have this data? Do we keep it simple and make data access the same to all teams and/or users or do we build in different access layers? Clearly there is a trade-off between acceptable risk and complexity.

To manage the different categories of data sensitivity, hoteliers should perform an audit and create a hierarchy of the categories of data, the requirement and use case and the consequence of loss or risk. Once this is created hoteliers can define the access policy and make acceptable risk versus complexity decisions.

Category	Example	Requirement / Use case		Consequence of loss / Risk	Access policy	Recommended action
PCI-DSS	Credit card number	Not needed		Significant / High	Strictly need to know	Delete at source
PII	Customer name	Not needed		Significant / High	Strictly need to know	Delete at source
Partner commercial	Booking value	By commercial teams for partner evaluation		Moderate	Commercial teams	Only available to commercial teams/users
Internal & partner commercial	Discount rate	By commercial teams for partner evaluation		Moderate	Commercial teams	Only available to internal commercial teams/users
Product information	Package names, list prices	Required by all for product ID		Limited	All internal users	Available to all teams/users

Figure 6: Data Access Audit

From a security perspective, data should be accessible only to specific individuals that are mapped to specific data security roles within the organization. One implementation scheme defines security roles and groups in an enterprise authentication tool such as Microsoft Active Directory, and all applications and storage systems in the enterprise inherit the data access rules that are setup and managed centrally in Active Directory. Special applications may require additional application rules to further restrict access.

Hoteliers should also consider that they may want contractors or other third-parties to access the data. In all cases an appropriate legal frame work – such as a non-disclosure agreement and data usage agreement – should be in place to protect and inform all parties. However, legal frameworks won't manage the reputational risk of data misuse so once more be conservative and only make sensitive data available to those that have a genuine need.

How Do Users Access the Data?

Allowing users to directly query the underlying, detailed data gives them flexibility and power. They can use whatever data they need and load the data into whatever tool is most beneficial to the use case at hand. In fact, for a data scientist who is building models, direct access to the data will be absolutely necessary. This is something to consider when evaluating third-party systems, as some off-the-shelf tools will not provide direct access to their underlying database or may provide only limited data export capabilities.

But for many users, direct access is not only unnecessary, it may actually be a drawback. The underlying model of a database is often complex and can easily be used incorrectly by a user who does not fully understand it. Something as simple as the difference between gross bookings and actualized ("paid-and-stayed") bookings can cause an incorrect analysis. In addition, not all users will have the technical skills to write queries that are correct and efficient in analyzing large amounts of data.

There are also many common calculations and aggregations that users will need to apply. Having these calculations and aggregations pre-built will keep users from making mistakes in writing them or applying inconsistent definitions, and will speed queries.

For these users, it can be very helpful to have pre-defined "data layers" that stand between the underlying database and the user. These data layers may be built in the database or may be built into the tools they are using to access the database (such as Tableau or Excel). These data layers can also help enforce data security. Data layers can be defined to only reveal aggregated data, hiding details like PII. In addition, different data layers can be created for different groups, so that each group only accesses the data they need to do their job.

Developing data layers does require an upfront investment, however. The hotelier must plan for what data, aggregations, and calculations will be needed to support the use cases and then implement the data layers in the database or tools. This is also likely to put additional daily processing requirements on the environment in order to refresh the data layers.

Off-the-shelf hospitality analytics tools will come with data layers already pre-defined. This way, the hotelier does not need to invest in building data layers itself. However, the data layers in these tools should be evaluated to ensure that they provide all the data access that the hotelier needs, as they may not allow the hotelier to define new data layers.

Similar considerations come into play when selecting tools that will connect to the database and used for analytics. You can allow users to select any tool that they want, which will give them maximum flexibility to use the tools they are comfortable with and most beneficial for the task at hand. For example, many users are still most comfortable working with data in Excel pivot tables, and this may not be possible if the hotelier is working with off-the-shelf tools that only allow predefined reports. In addition, data scientists certainly will require use of a language like R, Python, or SAS.

But allowing an unrestricted approach also has drawbacks:
- The more tools that are used, the more tools that need to be monitored and maintained to prevent security risks or attack vectors.
- Using different tools makes it more difficult for users to share commonly used queries, visualizations, models, etc., which creates rework and wasted effort.
- When users need to recreate their own versions of common queries etc. it increases the risk of inconsistencies, mistakes, and inefficient (or duplicative) queries and visualizations that increase environment load.
- It makes it harder for the company to establish a common look and feel for reports.
- Users may make mistakes in tools that they do not understand as well as they think, and there may be no one else at the firm who knows the tool well enough to catch the mistake.

Where Do Users Work With and Disseminate the Data?

Hoteliers should consider where users can work with the data. It may be easier for users if they are allowed to work with data directly on their workstations, whether in Excel, Tableau, Python, or another tool. Working in the data environment may be slower and certainly can put additional load on the servers. But duplicating data outside of the environment creates additional security risks, particularly if users are able to copy data onto USB drives or upload it to the Internet. To the greatest possible extent, hoteliers should limit data to a single, secured environment that is monitored and maintained.

The same factors apply not just to performing analytics but also to disseminating the reports and insights within the company. Hoteliers should consider how end users will access these reports – and what devices they will use to do so. For example will they view reports on phones, corporate mobile devices, home devices, or third-party devices? These are key consideration and need to be taken into consideration and reviewed with teams responsible for corporate information management policies and procedures. Obviously, hoteliers will need to strike an appropriate balance between security and accessing the data they need to run their business.

10 ANALYTICS SKILLSETS

Throughout the analytics planning process, it is important to consider what data and analytics skillsets you will have on your staff or available through vendor support.

This is obviously relevant to foundational infrastructure areas like data integration, analytics environment maintenance, and analytics tool selection. But it's also relevant to implementing your analytics use cases and to running them. For example, if you plan to have analytics run by your existing distribution team, then you will need to focus on self-service tools and simple use cases. Your existing employees probably do not have deep analytics skills. while it is possible to get them training, that will only go so far.

On the other hand, if you want to develop advanced, prescriptive models – like using machine learning to identify distribution patterns in massive amounts of data – then you are going to need some fairly advanced data science skills. At the least, you will need people who can integrate data, model data, select statistical models and program the models. This is a substantial commitment.

When deciding what kind of analytics skills you will need and how to acquire them, hotel companies should consider the many benefits of developing in-house capabilities. Early adopters are seeing a significant competitive advantage from distribution analytics, and this trend will continue. Companies that sit the sidelines with analytics will only fall further behind. Meanwhile, companies that outsource their analytics to a software vendor may catch up to their peers, but won't develop any proprietary advantage.

It is also worth remembering that analytics is an on-going process, not a one-time implementation. Many use cases will need to be run regularly, requiring analysis of the results each time. At the same time, since the hospitality market is constantly changing, there may also be a need to revise existing use cases and develop new ones. Anticipating these needs will be critical to determining your mix of analytics skills. For example, if you are building one complex model, and the rest of your frequent analytics use cases are fairly straightforward, you may want to outsource the model development while hiring for the recurring analytics.

In addition, companies often don't recognize that hiring analytics staff can actually increase retention amongst your other employees. Employees who get better data support from their company will make better decisions, achieve better results, and feel greater job satisfaction.

At the same time, there are a number of challenges that hotel companies face when building in-house capabilities. First of all, competition for analytics employees is fierce. That is true in all industries, but the pool of trained analytics resources is particularly small in hospitality.

In addition, it can be hard for hospitality companies to know who to hire. After all, if you don't have anyone doing analytics, who is qualified to vet candidates? Hiring blind is especially risky in analytics, because it is not easy to confirm the analysis results. Let's say that a data scientist writes a massive amount of R code to build a hierarchical grouped time series to predict future revenue trends. The results may look good, but who could validate that it was actually done correctly? This chicken-and-egg problem can be difficult to resolve.

Finally, hotel companies need to consider whether they can give analytics staff the professional and technical growth that they crave. People who work in analytics are naturally curious and eager to tackle new questions. In such a competitive labor market, they also expect rapid career progression. If you can't offer technical development and career progression, you may not be able to retain analytics employees very long, and you may be stuck in a cycle of hiring, training, and replacing.

For these reasons, some hotel companies will benefit by outsourcing to a third-party vendor that specializes in analytics. Hotel companies that go this route should seek a vendor that has an understanding of the hospitality industry. Technical skills and math knowledge aren't sufficient; to get meaningful insights, you need to understand the business domain. In fact, in a choice between advanced analytics skills and domain knowledge, in most cases the domain knowledge is the more important.

When outsourcing your analytics, you also need to consider other factors that are actually fairly similar to the ones you'd consider if hiring directly:

- Does the vendor have a good employee retention rate? If not, you may end up constantly retraining your vendor's resources in your business and in your data.
- Will your analyst(s) be dedicated and develop a deep understanding of your data and business? Or will you have a rotating stable of support?
- What proof can the vendor provide to show that its analyses are correct and create measurable value?
- Is there a cultural match between how the vendor approaches business and how your hotel company operates?
- How will the vendor ensure that your data stays secure while they are accessing it?

The Role of a Data Analyst

Many hotel companies will fall into the middle ground. They will keep analytics in-house, but hire individual data analysts rather than building an entire analytics organization.

The data analyst will have the analytics education and experience to derive insights from the data in support of business decision-making. In particular, the person will: work with the

business to understand what could be analyzed; prioritize the analyses; determine which data will be needed for the analysis; acquire the data; analyze the data; visualize the data; clearly communicate key findings; and make recommendations to improve business performance.

As you can see, the data analyst role requires a jack-of-all-trades. It is very unlikely that the data analyst will be able to build prescriptive models or apply advanced data science techniques; this requires specialists who would be part of a full analytics organization. But the data analyst can still derive insights from the data that the average employee could not.

Typically, the right candidate for a data analyst position possesses a natural curiosity. A data analyst is someone who enjoys figuring things outs, researches, asks the right questions, and has strong problem-solving skills. They want to know, not just because they should, or because someone told them to; they thrive on solving business problems with sound data.

When hiring a data analyst, or someone who will have responsibility for data analysis as part of his or her role, the following factors should be taken into account as part of the hiring process:

Professional Considerations:

- **Tools Experience:** What tools do they have experience of using: Are they familiar with the tools already installed, or similar systems. Are they familiar with which tools are suitable for which functions?
- **Creativity / problem solving:** While being able to follow house processes, are they equipped to deal with unplanned scenarios, which may require them to do things in a different way to get results. Can they suggest improvements?
- **Communication (strong verbal & written skills):** The more technical the person is, the more important this can be. Is the person capable of articulating the reasons why data is used to drive business decisions especially in an organization not currently driven by data decisions? Can this person explain technical or complex concepts to a non-technical audience? Conversely, can they take non-technical requirements from the business and translate them into mathematical analysis**?**

- **Ability to influence,** across the organization, within cross functional work groups and at the C level.

- **Attention to detail,** but also the ability to adapt and be flexible as requirements change.

- **Hospitality experience,** the ability to understand the industry without being confined to past thinking**.**

10. ANALYTICS TOOLS

Analytics tools will need to match business requirements and facilitate the transition from reactive to proactive decision making. There are a few avenues to be explored when considering the types of tools to implement. **Building your own** solution is a one of the foremost considerations for many hoteliers and hotel companies. And while it offers benefits such as full customization, complete control of data and flexibility; the initial development and maintenance costs along-side implementing the infrastructure for the visualization tools to support that, are most of the times outweighing the benefits. **Software as a Service** solutions on the other hand can offer advantages such as optimized user experience and increased data accessibility, low maintenance costs and continuous upgrades and enhancements. Another major benefit is the user training and familiarization. Some **Software as a Service** solutions work on data consumption quotas which is an important factor to be considered as the final cost could be much higher than initial estimations.

Analytics tools offer the opportunity to acquire a deeper understanding of the hotel business, its components and the factors that drive demand. They need to include a minimum array of features in order to provide with a flexible platform for the discovery, application and communication of insightful data and patterns. The end goal should be to empower users to retrieve, interrelate and visualize data with the purpose of offering a complete overview of the business and provide with key insights to drive smarter and more efficient decisions.

Analytics Tools: Attributes

While features need to match business requirements, the following represent a minimum set of attributes that will add value to the process of implementation:

- **Data integration:** to take full advantage data integration with current systems in place is imperative. Current systems can include your Property Management System, Distribution Systems, CRM's, Reporting tools etc.
- **Multi-Database:** a platform that supports multi-database integration will allow to easily access and retrieve data from various different sources
- **Built-in ETL (Extract, Transform & Load):** a built-in ETL capability will allow a seamless process of exporting data from a variety of systems, transform them into a single format in order for it to be imported into the data warehouse
- **Scalability:** as business needs change, the platform will need to allow for the constant development of customized insights to support latest trends and requirements
- **Data combination:** required for enriching insights through the combination of data from various sources
- **Intuitive user interface:** this will allow even the most novice user to get the most out of their interaction with the analytics platform while allowing a broader audience to use it
- **Easy to use & share**: an ideal application should not require specialized expertise to use, making it thus widely available to cross-departmental audiences. Sharing insights and ad-hoc or regular reports is an important attribute and the platform should allow defining different distribution groups and interval of distribution
- **Dashboards / Data visualization:** dashboards and data visualization are a vital element as they provide an effortless way to have a glance of key business metrics and bring data to life. As the human brain processes visual elements quicker than numbers; they make communicating and displaying performance indicators easy and effective
- **Drill down:** while dashboards provide with an easy way of keeping everyone in the loop, drill down features allow users to take a deeper and more exploratory approach and offer a broader and more detailed perspective of the data. Characteristics of such drill down features include filters, search functions, grouping, customized formulations etc.
- **Intelligent alerts:** this allows users to set predefined threshold restrictions that will allow a closer monitoring of how business is performing

- **Cloud compatibility:** cloud-based solutions make possible accessing and analyzing data business insights from anywhere and at any time. Data can be viewed from any device, (laptop, tablet, mobile etc.), enabling key personnel with the ability to make decisions on the go.

- **Data access:** data access enables retrieving information or raw data with the means of transferring to different systems or platforms

- **Mobile access:** with business levels constantly changing, access to the platform needs to be flexible for different devices such as smartphones, tablets etc.

- **Security:** with the rise in popularity of cloud-based solutions, security is an important factor to consider before migrating data to the cloud. It is imperative to validate that the solution is built and adheres to the latest security protocols and provides real-time back-up capabilities

Analytics Tools: Project requirements

Before setting out the implementation of analytics tools it is imperative to define the parameters, tasks and responsibilities that must be accomplished to ensure effective completion of the project. The purpose is to provide a clear picture of what the requirements are and the work that needs to be done in order fulfil the project through the alignment of resources and objectives. There are 3 main steps in defining the project requirements:

1. **Business requirements:** the aim is to set out project objectives and prioritize them by considering what information is required, how it will be used, by whom and what data transformation is required to reach that goal. Business requirements will need to be adapted to the company's structure and policies, while they secure the commitment from all relevant stakeholders, as execution will require changes in both set-up and operational procedures

2. **Solution requirements:** once business requirements have been outlined, it is important to consider the technical and configuration aspects for supporting the project. Based on the business and data needs that were specified in the previous step, the right platform will enable a smooth integration and configuration process while delivering the tools that are required.

At this stage it is important to consider the kind of tools to be implemented on top of the data. This can include generic BI tools that are used to analyze and visualize data and specialist tools such as revenue management systems, rate shopping tools etc.

3. **Stakeholders:** Communication is a key factor to successfully roll out and implement such a project. To achieve and maintain high levels of effective communication, it is imperative to highlight, right from the start, who the key stakeholders are and how they can influence, and be influenced, by the process.

For the higher levels of management this will include a detailed account of all aspects of the project and how implementation will affect current procedures alongside an on-going review of the progress. This will ensure their involvement and commitment.

For the rest of the team members a thorough introduction to working and maintaining the analytics tools will be a key factor to actively engage themselves while making the most of the systems capabilities.

11 CONCLUSION

An increasingly complex network of traditional and digital channels has to be managed to ensure hoteliers maximize their distribution to optimize occupancy, rates, and profitability. Data-based decisions are essential for driving the bottom line in the hospitality industry.

With a fixed capacity, a highly disposable product and high fixed costs, hotels are a natural candidate for making the most of the data that flows through their networks and systems. Distribution analytics can help hotels to better understand consumer buying patterns and the performance of their distribution strategies. Distribution data is important in supporting many other functions such as sales, marketing and revenue management to sell each room each night at an optimum price.

As part of the HEDNA Hotel Analytics Working Group, this document seeks to provide a core understanding of the fundamentals of distribution analytics which ties into the larger picture of revenue strategy. The document is structured to provide an insightful look into a whole range of different issues that should be taken into account when adopting and developing analytic capabilities across the organization. From defining what we mean by distribution analytics, to considering how to achieve essential data integration and data quality, and discussing thoughts on the skills, processes and tools needed, this paper is designed to offer some basic help and hands-on guidance for hotel companies both large and small. These investment considerations are fundamental and will empower hoteliers to make considered investments in the development of their analytic capabilities to meet the diverse and complex demands of today's competitive hospitality climate.