2022

Provincial Lists of Occupations in High Demand: Methodological Framework

Labour Market Intelligence research programme

higher education & training
Department:
Higher Education and Training
REPUBLIC OF SOUTH AFRICA

NSF
National Skills Fund
FUNDING TO SKILL OUR NATION

LABOUR MARKET iNTELLIGENCE

DPRU
DEVELOPMENT POLICY RESEARCH UNIT

© Published in 2022 by:

Department of Higher Education and Training
Private Bag X174
Tshwane, 0001
www.dhet.gov.za

ISBN: 978-1-77018-940-9

2022

Provincial Lists of Occupations in High Demand: Methodological Framework

Labour Market Intelligence
research programme

Authors:
Fouché Venter, Michele Capazario

Acknowledgements

The authors would like to thank everyone who assisted with and contributed to this report. We extend our gratitude to the Department of Higher Education and Training (DHET) and the Development Policy Research Unit (DPRU) for their guidance and input throughout the process. In particular, the contributions of Ms Mamphokhu Khuluvhe (DHET) and Dr Hersheela Narsee (peer reviewer) were invaluable to the process of developing the methodology.

Suggested citation

Venter, F.H.J., Capazario, M. (2022). *Provincial Lists of Occupations in High Demand: A Methodological Framework*. Labour Market Intelligence research programme, Department of Higher Education and Training.

Table of contents

List of figures

List of tables

Abbreviations

ARIMA	Autoregressive Integrated Moving Average
DHET	Department of Higher Education and Training
GVA	Gross Value Added
LMDS	Labour Market Dynamics Survey
NQF	National Qualifications Framework
OFO	Organising Framework for Occupations
OIHD	Occupations in High Demand
PCA	Principal Components Analysis
P-OIHD	Provincial Occupations in High Demand
PSET	Post-School Education and Training
QLFS	Quarterly Labour Force Survey
SASCO	South African Standard Classification of Occupations
Stats SA	Statistics South Africa

PART 1

Introduction

The first national list of Occupations in High Demand (OIHD) was published by the Department of Higher Education and Training (DHET) in 2014. Since then, a new iteration of the list has been produced every two years, each incorporating new developments in the country's labour market.

The national list of OIHD is an important skills planning mechanism. Its primary purpose is to inform planning in the Post-School Education and Training (PSET) sector by:

- Signalling the need for new qualifications;
- Acting as a signpost for enrolment planning;
- Guiding resource allocations; and
- Informing career guidance for learners and work-seekers. (Capazario & Venter, 2020)

To achieve these objectives, the national list of OIHD identifies occupations across the economy that meet the criteria shown in Figure 1.

FIGURE 1: The concept of an occupation being in high demand

Show/could show relatively strong signs of employment growth

Are in shortage (meaning that the demand for individuals within an occupation is higher than the supply of adequately and appropriately skilled individuals available in the labour market)

Are expected to emerge as being in demand in the near future as a result of technological advancements and innovation, or government's strategic priority

Source: Adapted directly from Reddy, Rogan, Mncwango and Chabane (2018)

Several provinces have made calls for provincial lists of OIHD (P-OIHD) during meetings held by the provincial skills development fora. Their motivation is clear: Without province-specific research, provinces use national occupational demand to infer provincial demand. However, provincial demand often deviates from national demand. Given different provincial contexts, occupations in high demand in one province might not be in high demand in another.

For this reason, the express purpose of this document is to propose a research methodology that can be applied ubiquitously across provinces. The P-OIHD methodology is adapted from the national OIHD methodology but exclusively uses provincial-level data and inputs from province-specific stakeholders.

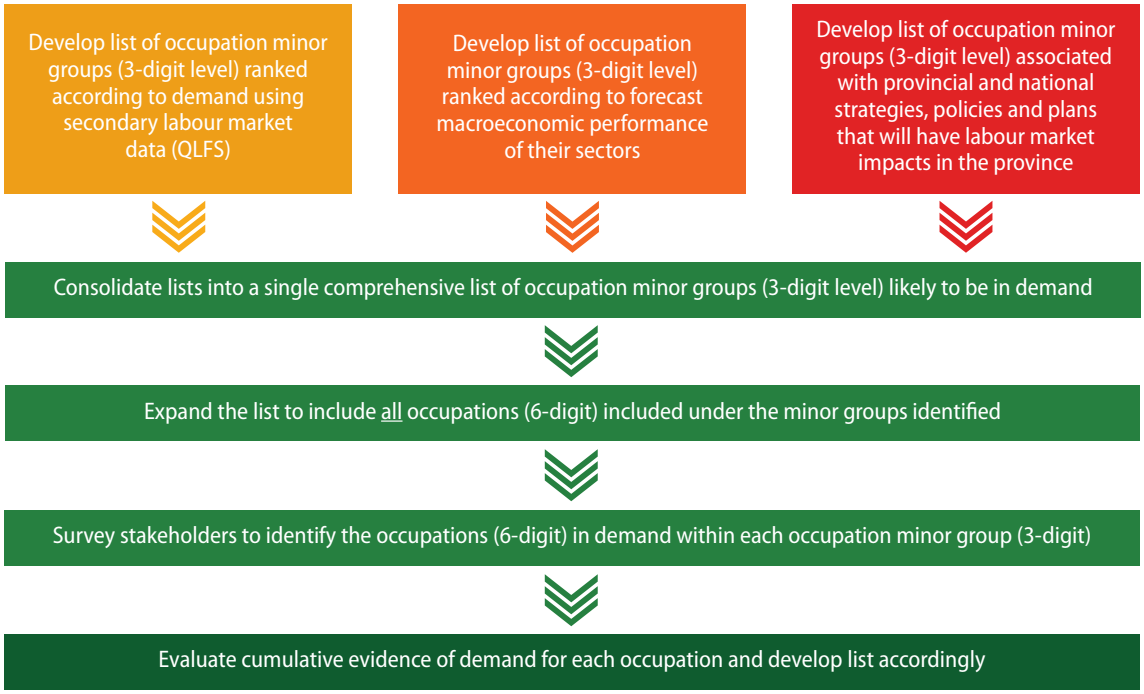
PART 2

Methodology

South Africa's Organising Framework for Occupations (OFO) is an essential building block of the methodology. Therefore, familiarity with the framework is an important precursor to reviewing and applying the methodology. The central components of the OFO are discussed and explained in Appendix 2 of this report.

The following is a simple process flow diagram representing the proposed approach to developing a list of P-OIHD. The importance of the OFO is immediately apparent.

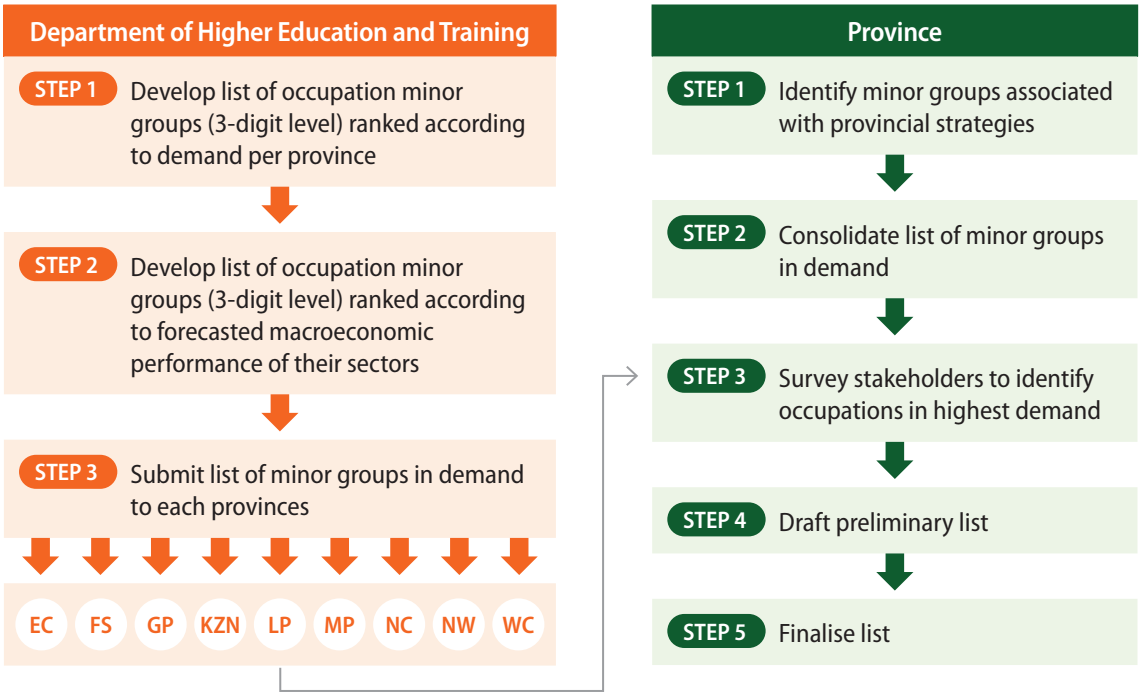
FIGURE 2: Process flow to construct a list of P-OIHD



Although the development of each list of P-OIHD is the province's responsibility, the DHET has an important role to play. Because the quantitative analysis can be done for all nine provinces simultaneously, it requires substantially less time and effort when done centrally.

Figure 3 expands on the steps set out in Figure 2 and proposes a split allocation of responsibility between the province and the DHET to minimise the time and effort required to develop the lists and ensure alignment between provinces.

FIGURE 3: High-level methodological overview



The three steps required from the DHET and the five steps required from the province are described in Sections 2.1 and 2.2, respectively.

2.1 Steps for the DHET

Before setting out the quantitative analysis, it is important to note that it is sophisticated and will require advanced expertise in statistical analysis. Step 2 will specifically require an econometrician for the development of economic forecasts.

2.1.1 Step 1: Rank minor groups (3-digit) according to demand per province

The ranking of minor groups (at the 3-digit OFO level) is done by simultaneously considering multiple demand indicators. These indicators aim to assess employment pressure, wage pressure and vacancy pressure – all indicators of increasing demand. Table 1 provides detail on the rationale for using each indicator; how each indicator is calculated; the indicators' relationship to occupational demand; and the data source used for the assessment of each indicator.

TABLE 1: Indicators of demand at the provincial level

DIMENSION	RATIONALE	SOURCE	VARIABLE (INDICATOR)	RELATIONSHIP WITH DEMAND
Employment pressure	When analysed simultaneously, an increase in the number of people employed (growth), an increase in the number of hours worked (intensity), or a decrease in the duration of employment tenure (duration) is associated with increased occupational demand.	Quarterly Labour Force Survey	Employment growth	Positive
			Employment intensity	Positive
			Employment duration	Negative
Wage pressure	Upward pressure on wages could signal that employers are willing to pay more for labour than they did previously, making the case that their demand for labour has increased relative to the supply of that labour.	Labour Market Dynamics Survey	Mean wage growth	Positive
			Median wage growth	Positive
			Conditional mean wage growth	Positive
Vacancy pressure	An increase in vacancies or the average advertisement period signals that demand is not being met.	Web-based employment services[1]	Vacancy growth	Positive
			Vacancy duration	Positive
			Supply–demand gap growth	Positive

Once the value of each variable has been calculated, the following steps should be followed:

1. Rank minor groups according to each indicator
2. Calculate the average ranking per dimension
3. Calculate the average demand ranking for each minor group (m) using the following:

$$\text{Average Demand Rank}_m = \frac{\text{Average Employment Pressure Rank}_m + \text{Average Wage Pressure Rank}_m + \text{Average Vacancy Pressure Rank}_m}{3}$$

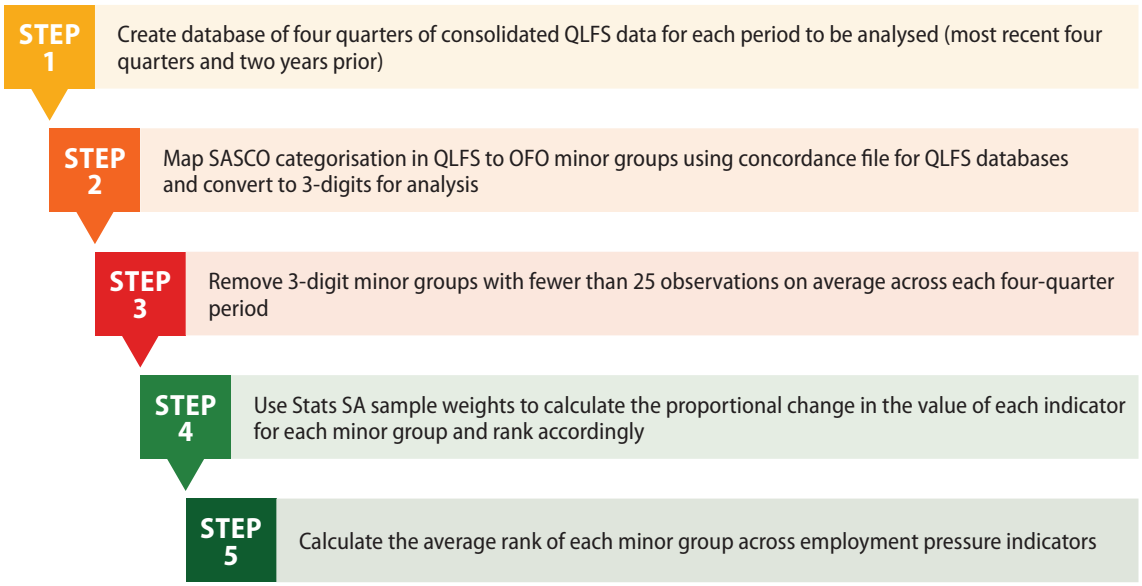
Calculating each dimension's average rank (Step 2) involves some nuance. Therefore, each sub-step is described separately in the next section.

1 Career Junction was used for the national list of OIHD.

Employment pressure

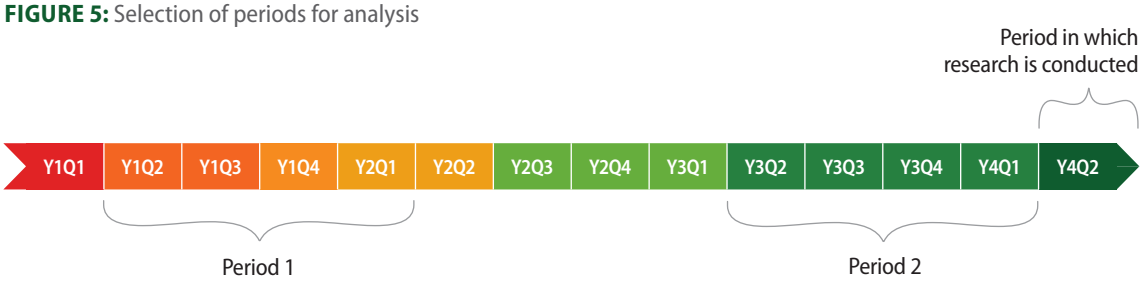
The following process is applied for each employment pressure indicator using data from the Quarterly Labour Force Survey (QLFS):

FIGURE 4: Employment pressure process map



1 Consolidate four quarters of the QLFS for the two time periods between which growth will be measured.

FIGURE 5: Selection of periods for analysis



2 Map the SASCO codes to the OFO.

The QLFS includes the current occupation of each of the employed survey participants. However, Statistics South Africa (Stats SA) uses the South African Standardised Classification of Occupations (SASCO) codes to classify occupations. Therefore, it is necessary to reclassify each participant's occupation according to the OFO.

A concordance file is available from the DHET, which allows for the automatic completion of the reclassification. It is important to note that the concordance file reclassifies the occupations at the 4-digit level. Therefore, the last digit of each 4-digit code can be removed as the analysis will be done at the 3-digit level.

3 Remove 3-digit minor groups with fewer than 25 observations on average across the four quarters.

This exclusion prevents small-sample bias, which could potentially skew results. Importantly, the exclusions will not be final. Instead, if respondents to the surveys (Step 3 of the provincial steps) provide strong, evidence-based representation to include occupations from these excluded minor groups, these occupations should be reincluded.

4 Use Stats SA sample weights to calculate the proportional change in the value of each indicator for each minor group and rank accordingly.

Stats SA assigns sample weights for each observation in the QLFS. In a rudimentary sense, these weights represent the number of individuals in the South African population who would provide similar responses to the survey that the respondents did.

The weights are incorporated into the calculation as follows:

- Employment growth for a minor group = proportional change in the sum of weights of a minor group from period t–2 to period t
- Employment intensity change for a minor group = proportional change in the weighted average number of hours worked per observation from period t–2 to period t
- Employment duration change for a minor group = proportional change in the weighted average duration of employment per observation from period t–2 to period t

5 Calculate the average rank of each minor group across employment pressure indicators.

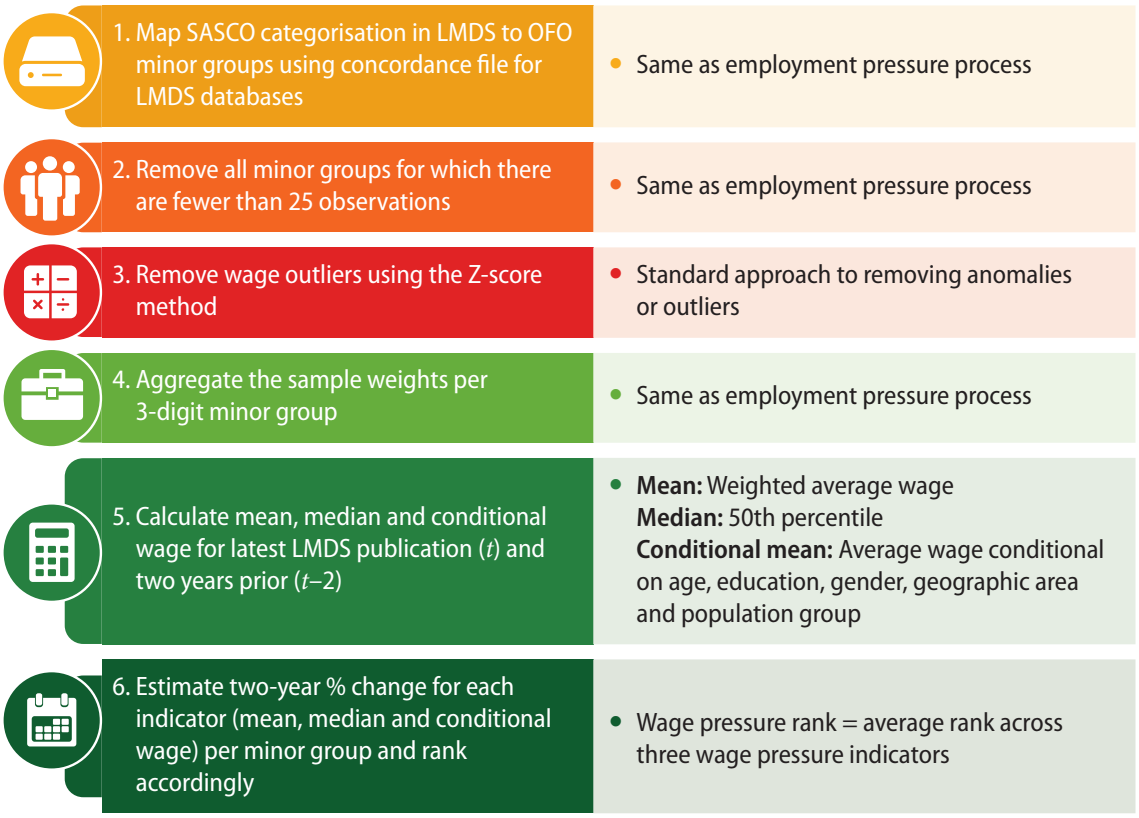
Once the indicator values have been calculated, the minor groups are ranked separately for each employment pressure indicator. These ranks are averaged to provide an average employment pressure rank for each minor group as follows:

$$\text{Average Employment Pressure Rank}_m = \frac{\text{Employment Growth Rank}_m + \text{Employment Intensity Rank}_m + \text{Employment Duration Rank}_m}{3}$$

Wage pressure

The process map below summarises the steps required for analysing wage pressure using the Labour Market Dynamics Survey (LMDS).

FIGURE 6: Wage pressure process map



The LMDS contains information on wages and salaries. As before, matching the SASCO codes in the LMDS to OFO codes is crucial for framing the analysis. This process is also followed by the exclusion of minor groups with less than 25 observations from the analysis (and, once again, developing a list of minor groups excluded as a result, which then requires validation by the provinces).

The removal of outliers in the data is proposed due to their tendency to skew results. The following process can be followed (Step 3):

1. Create a monthly wage variable by adding two variables already found in the LMDS:
   - Q54a_monthly (only if Q54a_monthly is left blank)
   - Q57a_monthly (only if Q54a_monthly is left blank)
2. Where both Q54a and Q57a are blank, the monthly wage can be estimated as the midpoint of each income bracket found in Q58SALARYCATEGORY
3. Calculate average monthly wage per minor group
4. Calculate the monthly wage standard deviation per minor group

5. Calculate Z-score per observation:[2]

$$\frac{\text{Observation monthly wage}_m - \text{Average monthly wage}_m}{\text{Monthly wage standard deviation}_m}$$

6. Remove observations where Z score > |3|

(Sharma, 2021)

Once all outliers have been removed, the wage pressure indicators can be calculated. Calculating the growth rate in the mean and median wage over the two years is straightforward. The most important thing to remember is the sample weights. For the calculation of the mean wage, one has to calculate the weighted average wage for each period:[3]

$$\text{Weighted average wage}_m = \frac{\Sigma_n \text{weight}_{mn}\ \text{wage}_{mn}}{\Sigma_n \text{weight}_{mn}}$$

In the case of median wage growth, the below steps are required for each minor group. (A hypothetical example is provided for clarity.)

1. Rank observations from highest to lowest

WAGE	WEIGHT ASSIGNED TO OBSERVATION IN LMDS	CUMULATIVE WEIGHT
R43 559	30	30
R40 872	20	50 = 30 + 20
R39 250	43	93 = 43 + 20 + 30
R36 053	49	142
R32 726	14	156
R31 840	23	179
R24 196	14	193
R20 351	44	237
R18 583	18	255
R9 438	15	270
R8 489	42	312
R3 850	49	361

2. Calculate the cumulative weight for each observation: add weights of all observations up to and including the observation
   - See Column 3 above

2 *m* represents the minor group.
3 *m* represents the minor group and *n* the number of observations in the data set for minor group *i*.

3. Halve the aggregate weight of all observations
   - Aggregate weight of all observations is 361
   - Halve = 180.5
4. Identify the median: the monthly wage of the observation that corresponds to the number calculated in Step 3
   - Closest cumulative weight to 180.5 is 179, which corresponds to a wage of R31 840, which is the median
5. Calculate the % difference between this median wage and each minor group's monthly wage

The calculation of the conditional wage is slightly more complex. Conditional wages are wages conditional on sociodemographic factors such as age, gender, race and geographical location. This indicator is critical to include in the analysis because it controls for factors that would influence wages besides increasing demand.

The following regression is used to estimate the conditional wage per minor group:[4]

$$\log(\text{monthlywage}) = \beta_1 \text{MG} + \beta_2 \text{Province} + \beta_3 \log(\text{Q14AGE}) + \beta_4 \text{Q17EDUCATION} + \beta_5 \text{Geotype} + \beta_6 \text{Q13GENDER} + \beta_7 \text{Q15POPULATION}$$

Estimating this model (i.e., estimating the **β's**) requires the following coding steps in Stata:

1. **svyset uqno [pweight = Weight]** /* Ensures that the regression considers the weights assigned to each observation*\
2. **svy: regress log(monthlywage) i.MG i.Province log(Q14AGE) i.Education i.Geo_type i.Q13GENDER i.Q15POPULATION** /* MG = minor group, Geo_type = rural traditional/rural farms/urban, *i* tells Stata to create dummy matrix for variable*\

3. Once the coefficients have been estimated, a conditional wage variable, *condwage*, should be generated. The *condwage* for each observation is calculated using the coefficients estimated by the regression model. Finally, a weighted average *condwage* per minor group (m) should be calculated:[5]

$$\text{Weighted average condwage}_m = \frac{\Sigma_n \text{weight}_{mn}\ \text{condwage}_{mn}}{\Sigma_n \text{weight}_{mn}}$$

4. This conditional wage analysis is done for periods t and $t-2$, and the growth between the two periods is calculated.

4 This model is estimated separately for each period's data set (t and $t-2$).
5 n represents the number of observations for minor group m.

5. The final step is to rank the minor groups according to each wage pressure indicator as was done with employment pressure. These ranks are averaged out to provide an average wage pressure rank for each minor group.

$$\text{Average Wage Pressure Rank}_m = \frac{\text{Mean Wage Growth Rank}_m + \text{Median Wage Growth Rank}_m + \text{Conditional Wage Growth Rank}_m}{3}$$

Vacancy pressure

This data is not available publicly. Instead, data should be purchased directly from an employment services provider (Career Junction has been used in the past). Importantly, this data will automatically match not the country's OFO codes but the service provider's nomenclature. As a result, parties requesting this data should provide the service provider with the OFO or map the OFO to the service provider's nomenclature and job titles before requesting vacancy data. Either option has cost and time implications, but historically the former process has been pursued in analysing similar lists.

The following data should be requested:

- The number of vacancies per minor group for periods t and $t-2$;
- The number of vacancy renewals per minor group for periods t and $t-2$; and
- The number of applications per minor group for periods t and $t-2$.

Once received, the vacancy pressure indicators are calculated by determining the percentage difference between period t and $t-2$ for the number of vacancies, the number of vacancy renewals, and the supply–demand gap (SDG).[6] As with the other dimensions, the minor groups are then ranked, and an average vacancy pressure indicator is estimated:

$$\text{Average Vacancy Pressure Rank}_m = \frac{\text{Vacancy Growth Rank}_m + \text{Vacancy Duration Rank}_m + \text{SDG Rank}_m}{3}$$

With this average ranking complete, the average ranking per dimension can now be used in the average demand rank formula.

2.1.2 Step 2: Rank occupations according to macroeconomic prospects

The analysis in Phase 1 considers recent employment trends and infers potential future dynamics. Although these trends have meaningful predictive value, other economic factors also play a significant role in the labour market. An industry's economic performance impacts the number of individuals that industry will employ;[7] as an industry grows, so does its demand for individuals to work within it. Therefore, Step 2 considers the economic prospects of industries and how these influence future occupational demand.

6 SDG equals the difference between applications (supply) and vacancies (demand).

7 See Aguiar-Conraria, Martins and Soares, 2019; Ball, Furceri, Leigh and Loungani, 2019; Onakoya and Seyingbo, 2020; Gonzalez and Jontxen, 2020.

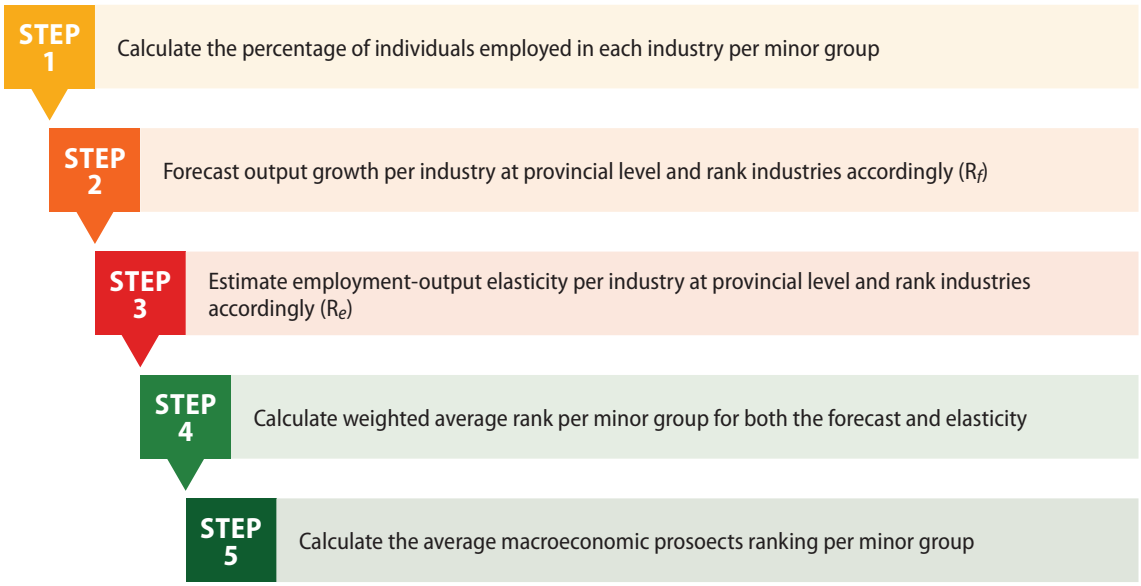
Two indicators are considered for each industry, as shown in Table 2.

TABLE 2: Indicators of macroeconomic prospects at the provincial level

DIMENSION	VARIABLE	DESCRIPTION
Macroeconomic prospects	Gross Value Added (GVA) growth	Forecasted percentage change in GVA over next two years
	Employment-output elasticity	The historical relationship between GVA and employment

The process shown in Figure 7 is followed to rank minor groups according to their respective economic prospects.

FIGURE 7: Macroeconomic prospect process map



The first three steps in Figure 7 do not have to be done in any specific order as they are independent. However, the order presented in the figure is how this process will be described in this section.

Each observation in the QLFS includes the industry in which the respondent is employed. Across the same minor group, multiple industries could be listed, as occupations are often salient across multiple sectors. Even when analysing the industry distribution at the occupation (6-digit) level, occupations are often required across multiple industries.

If we consider future occupational demand based on forecasted industry performance, we need to understand how employment is distributed across industries. To approximate this distribution, we calculate the distribution of observations (based on their survey weights) across the industries for each minor group.

$$D_{im} = \frac{\Sigma_i \text{weights}_{mi}}{\Sigma_m \text{weights}_{mi}}$$

D_{im} = Percentage of minor group m in industry i

This industry-level breakdown is done at the highest aggregate industry classification possible, using the *indus* variable in the QLFS. The *indus* variable classifies minor groups into the industries shown in Figure 8.

FIGURE 8: Industry classifications in QLFS and LMDS

Although not as granular as would be ideal, this high degree of aggregation is necessary to ensure that the data is reliable at the provincial level.[8]

The two remaining steps in Figure 7 require econometric analysis. Stats SA publishes annual industry-level economic output data for each province. This data can be used to forecast[9] each industry's expected output, which, when combined with QLFS data, can determine the trajectory of the provincial macroeconomy and labour market absorption potential.

8 The data in the QLFS/LMDS is representative at the national level. Further disaggregating by province weakens the reliability of the data to a degree. Further slicing the data by industry and occupation stretches the reliability of the data even more. Therefore, although the QLFS/LMDS has more granular industry classification variables, a further industry disaggregation would make the data almost wholly unreliable and prone to inaccuracies.

9 It is important to flag here that point forecasts are, more often than not, imprecise. However, given the context of the data and how it is being used, what forecasts are relatively robust at providing is a sense of the trend of particular data. It is predicting either an upward or downward trend for data that inevitably assists in assessing a 3-digit minor group's macroeconomic prospects.

Gross Value Added (GVA) indicates the former, and employment-output elasticity indicates the latter (i.e., labour market absorption potential). These concepts are discussed in more detail next.

1 Forecasting GVA:

- GVA measures the contribution to GDP made by an individual producer, industry or sector over a specified period (OECD, 2021).
- Although the researcher could use any preferred but defensible forecasting technique, this analysis is suggested to be kept simple.
- In Stats SA's fourth-quarter publication of Gross Domestic Product, GVA is provided for each industry within each province.

The authors propose using an Auto-Regressive Integrated Moving Average (ARIMA) model for forecasting GVA, given its ease of implementation relative to alternative approaches.[10]

Essentially, an ARIMA model uses past variations over time to forecast. Employing an ARIMA model requires using the Box-Jenkins to specify the model. For this application, and given its user-friendliness, the authors propose utilising EViews to conduct the ARIMA modelling (although this can also be done in a similar vein using Stata). A step-by-step guide to doing this is presented in Annexure 2.

2 Estimating employment-output elasticity:

- Employment-output elasticity indicates how well economic growth translates into job creation. Put differently, employment elasticity indicates how much employment in an industry would grow if industry-level GVA grew by 1%.

6. The coefficient (β_1) in the following OLS regression is the employment-output elasticity of industry i and province p:

$$\log(\text{employed}_{ip}) = \beta_1 \log(\text{GVA}_{ip})$$

The indicators should be considered together. A similar ranking approach to the one in Phase 1 should be taken:

1. Rank industries according to GVA prospects.
2. Rank industries according to employment-output elasticity.
3. Calculate weighted ranking of minor groups using D_{im}:

$$\text{Weighted ranking}_{im} = \sum_{i} \text{Rank}_{im} \times \text{D}_{im}$$

4. Rank minor groups according to weighted ranking.

10 Should researchers have capacity to conduct more robust forecast analysis using alternative methods (for example, based on social accounting matrices, computable general equilibrium models or panel vector autoregression models), such a forecasting method should also be considered. Given the relative simplicity and ubiquity of employing an ARIMA, however, this model specification has been suggested.

2.1.3 Step 3: Submit list of minor groups in demand to each province

Two ranked lists of minor groups will be available once Steps 1 and 2 are completed. The first ranks minor groups according to a combined consideration of employment, wage and vacancy pressure (Step 1). The other list ranks minor groups according to their economic prospects (Step 2).

These rankings will be subject to a cut-off to shortlist the minor groups in the highest demand. In line with the cut-off used in the national list of OIHD, the following should be used:

- The top 25% of the minor groups ranked according to employment, wage and vacancy pressure; and
- The top 25% of the minor groups ranked according to economic prospects, provided they are at least in the top 40% of the minor groups according to employment, wage and vacancy pressure.

It bears mentioning that the choice of cut-off is based more on historical precedent than statistical analysis. The 2018 and 2020 iterations of the national list of OIHD utilised the 25% cut-off to signify whether a minor group/occupation showed sufficient demand evidence.

On the other hand, very little analysis in South Africa has been done relating macroeconomic progress to labour demand at the minor group level. As a result, requiring a minor group to rank among the top 40%, as stipulated in Point 2, is not based on theory per se. These cut-offs can be debated, but ultimately, very little literature exists that guides exactly how high such a cut-off should be, making the choices at least somewhat subjective.

The subjectivity of this cut-off point is one of the reasons the qualitative steps performed by the process are so important. Beyond just the cut-offs, the quantitative steps described above are somewhat rigid, and it is important to incorporate flexibility through the survey process. Therefore, even if an occupation is excluded from these preliminary lists, it can still make it onto the final list if the survey process offers sufficient evidence.

2.2 Steps for the province

2.2.1 Step 1: Identify minor groups associated with provincial strategies

The provinces' first step aims to identify minor groups that might experience a surge in demand due to government strategies. Provinces do not have to wait for the quantitative analysis to be completed to commence this review as the two processes are independent.

Government strategies rarely specify the skills or occupations for which demand will increase. Therefore, it is suggested that a value chain analysis be used to identify the products (goods and services) for which demand and supply will increase due to strategy implementation. The identified products should then be used to identify minor groups in the OFO. At a high level, this mapping and value chain exercise will move along the process flow shown in Figure 9.

FIGURE 9: Process flow of moving from government strategy to identifying minor groups likely to experience increased demand

The following procedure is suggested:

1. Researchers should request that all provincial and municipal[11] entities submit strategies or plans that are being implemented or that will be implemented over the coming 12 months;
2. Researchers should then identify products (goods and services) that will see expansion as a result of the implementation of these strategies;
3. This identification process will assist in producing a high-level description of the value chain for each of the products identified;
4. After developing a value chain overview, researchers should identify keywords that drive the value chain activities;
5. These keywords should be used as search terms to identify minor groups pertinent to each identified value chain from OFO descriptors;[12] and
6. The findings from the identification of minor groups should be sent to the custodian of the strategy or plan for validation and confirmation.

Box 1 Value chain analysis: Suggestions

It is crucial to flag that value chain and keyword analysis are not necessarily standardisable to the smallest detail. However, the authors have conducted similar value chain analyses before and therefore offer some key facts to simplify the process.

- First and foremost, multiple value chain analyses already exist for certain products and services.
- If there is no value chain available in the literature for the product or service identified, the researchers' next step would be to develop a value chain overview diagram. This diagram should summarise all possible inputs into and outputs from the product or service.
- Once such a diagram is developed (or adapted from the literature), it should be used as the basis for the keyword search. For example, if a good requires transportation, keywords would include all types of transportation applicable (air/aeroplane/drone, water/boat/ship, road/car/truck/van, etc.) as well as any terms linked to the transportation of goods (logistics, safety). Each product or service identified should have a keyword database of no more than 10 words, which will then be cross-checked with occupational descriptions in the OFO. All matching descriptors will be boiled down to their 3-digit components, and flagged as matching the strategic priority criterion.

11 The objective of the district development model is to foster intergovernmental relations, and further ensure that the distance between national, provincial and local government – on the one hand – and their constituencies – on the other – is narrowed. By including municipalities, labour market challenges can be addressed more systematically, with municipal level issues that were not focused on previously now coming to the foreground (Department of Corporate Governance and Traditional Affairs, 2020).

12 This step requires an in-depth knowledge of the OFO. The key words identified should match the descriptors in the OFO, which is time-consuming, but ultimately more accurate than simply looking at the minor group titles.

2.2.2 Step 2: Consolidate lists of minor groups in demand

Once Step 1 is concluded, the researchers will consolidate the following list of minor groups:

- The DHET will have provided the following:
  - The top 25% of the minor groups ranked according to employment, wage and vacancy pressure; and
  - The top 25% of the minor groups ranked according to economic prospects, provided they are at least in the top 40% of the minor groups ranked according to employment, wage and vacancy pressure.
- The province will have compiled the following:
  - Minor groups identified as likely to experience increased demand as a result of provincial strategies.

This information, once consolidated, forms the basis for the survey to stakeholders.

2.2.3 Step 3: Survey stakeholders to identify occupations in the highest demand

While the list of minor groups is useful at this stage, the level of detail is insufficient. For this reason, surveys should be used to move the list from a set of 3-digit minor groups to more granular 6-digit occupations. The demand identified at the 3-digit level may be driven by only a portion of the 6-digit occupations within the category. The survey aims to identify those occupations.

Surveys should be sent to provincial government departments, municipalities, SETAs, industry bodies and employers.

It will be particularly important to ensure that the survey covers employers across both the industry and the size spectrum. The Department of Economic Development's relationship with employers could be leveraged to reach a large proportion of the desired coverage. However, reaching micro and small enterprises may be a challenge. For these SMMEs, municipalities may be the best place to disseminate the survey. Although online surveys should be the primary approach to dissemination, telephonic and in-person surveys are also encouraged if affordable.

With sufficient sectoral participation, evidence from such a survey will provide a balanced view of both private and public sector institutions, reflecting government strategic priorities and the needs of the private sector alike.

The survey should be extremely straightforward to complete and should follow the ideal of maximising response rate by minimising response time. Two suggestions, which have seen success in past iterations of occupational lists, are offered below.

1 Utilise a Likert scale.

Proposed Survey Question 1

The following minor group (3-digit) occupations were identified as being in high demand in [province]. It is, however, unlikely that all 6-digit occupations within this category are indeed in high demand. Indicate, with an "x", your opinion regarding the demand for these matched 6-digit occupations over the next two years in [province]:

MINOR GROUP	LINKED OCCUPATIONS	SCORE = 1 (LOW DEMAND FOR OCCUPATION)	SCORE = 2 (MODERATE DEMAND FOR OCCUPATION)	SCORE = 3 (HIGH DEMAND FOR OCCUPATION)	NOT APPLICABLE TO MY INSTITUTION
2019-xxx	2019-xxx111		x		
	2019-xxx112			x	
	2019-xxx113				x
	...				

Proposed Survey Question 2

Please list occupations not included in the previous question that you think will experience high demand over the next two years in the space provided. Also indicate the degree to which they will be in demand, and the reason:

PROPOSED OCCUPATION	SCORE = 0 (NO DEMAND FOR OCCUPATION)	SCORE = 1 (LOW DEMAND FOR OCCUPATION)	SCORE = 2 (MODERATE DEMAND FOR OCCUPATION)	SCORE = 3 (HIGH DEMAND FOR OCCUPATION)	REASON/ EVIDENCE
Welder			x		
Electrician				x	
Plumber			x		
...					

The following types of evidence, among others, can be classified as sufficient:

- Documentation identifying large investment or expansion plans;
- Government strategic documents; and/or
- Representative surveys indicating high demand or shortage.

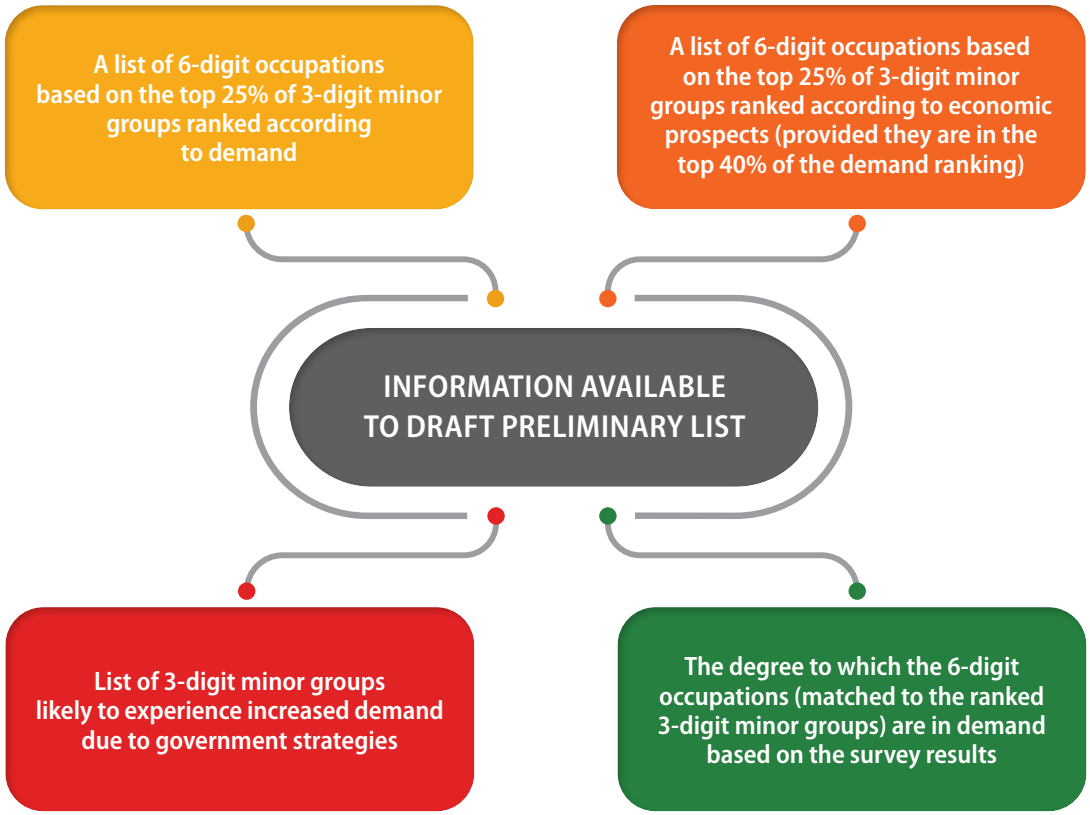
2 Target minor group survey at industries.

The list of minor groups should only be sent to stakeholders in industries relevant to the minor group. For example, an employer in the finance industry should not be asked to indicate its view on the provincial demand for veterinarians. An industry is considered relevant to an occupation if more than 10% of the individuals employed within a particular minor group (according to aggregated QLFS for the latest period) fall within that industry ($D_{im}>10\%$).

2.2.4 Step 4: Draft preliminary list

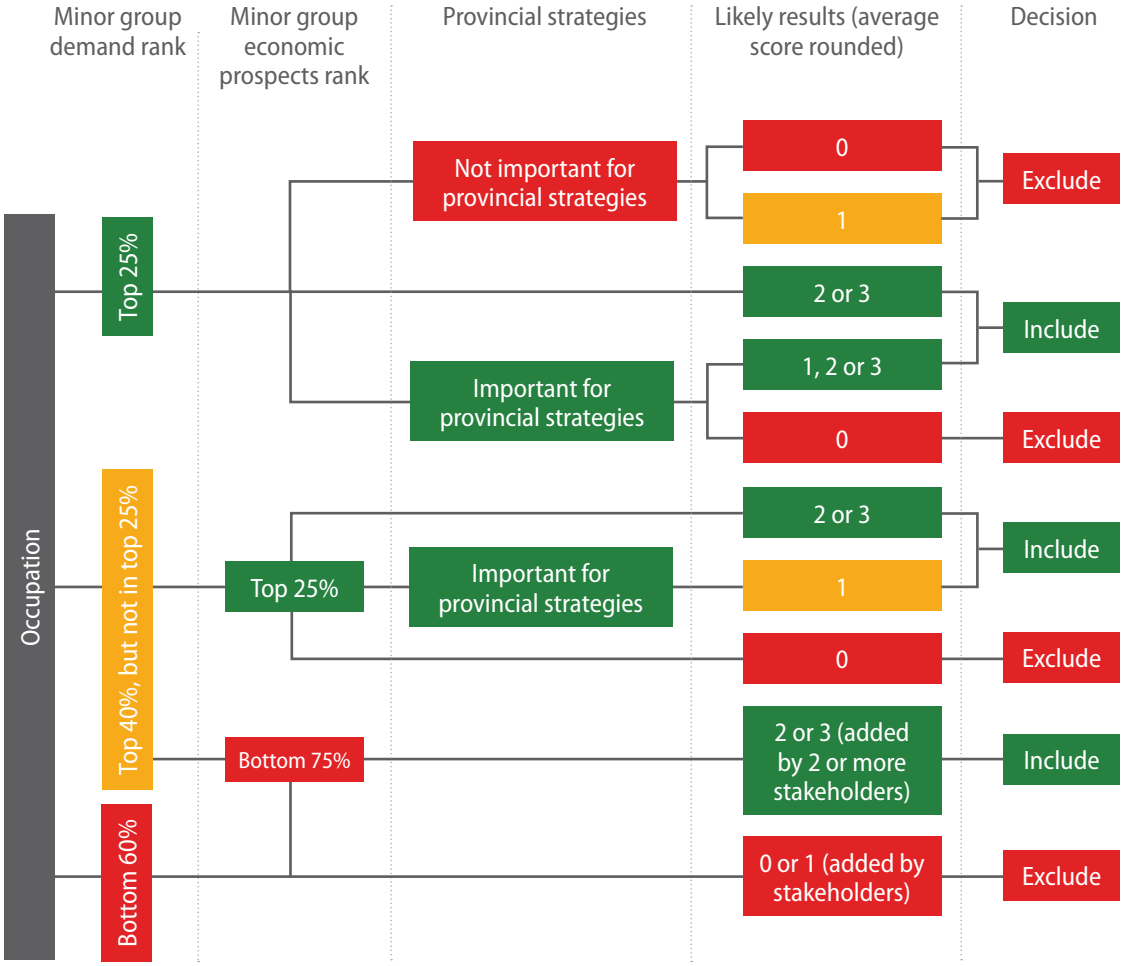
With Step 3 complete, the information shown in Figure 10 is now available.

FIGURE 10: Information required to draft a preliminary list



The evidence, described in Figure 10, can now be used to apply exclusion/inclusion criteria to determine the preliminary list ready for validation. The criteria can be applied through the algorithm presented in Figure 11.

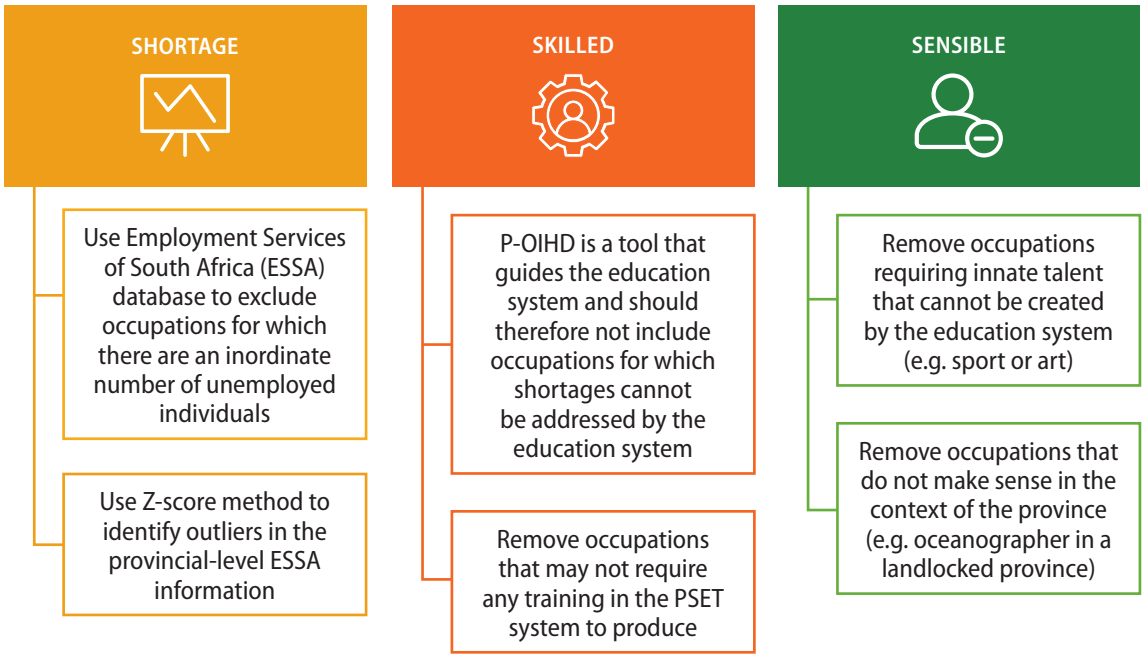
FIGURE 11: Inclusion and exclusion criteria for the list of P-OIHD for public comment



2.2.5 Step 5: Finalisation of list

Once the algorithm is applied, a preliminary list will be ready for finalisation. The finalisation should be done through an internal validation process. Three factors should be considered **in tandem** as part of this process (see Figure 12).

FIGURE 12: Final validation process of the list of P-OIHD (3S method)[13]



Source: The Migration Advisory Council (2013)

It bears mentioning here that there is also room for the national list of OIHD to feed into the list of P-OIHD for each province. Where there is insufficient evidence provided by stakeholders or where stakeholders provide provinces with evidence that is at odds with conventional labour market wisdom, the national list of OIHD can be used as a sensibility filter. Occupations on the periphery of being excluded should be reincluded into the list of P-OIHD if the occupation is found on the national list of OIHD and is sensible for the province to include (i.e., an oceanographer would not be sensible to include for a landlocked province).

It is expected that the national list of OIHD will be amended/updated with inputs[14] from each P-OIHD submission, ensuring that the list of P-OIHD feeds into the national list of OIHD and subsequently that the latter can be used for validation of the former.

13 It is important to flag that the national list of OIHD excluded all elementary occupations under the "skilled" sense check. However, in the case of the list of P-OIHD, the researchers feel it is important to include elementary occupations where such occupations require training in the PSET system. "Elementary" does not necessarily refer to occupations that do not require any form of training, but there is also a subset of elementary occupations for which there may be no PSET training needed, with the latter being excludable from the list.

14 This reciprocity extends both to those occupations that may not appear in the national list but that appear in the provincial lists, and to the advancements in the P-OIHD methodology.

PART 3

Conclusion

This report has described the methodology used to draft provincial lists of OIHD in South Africa. As with the national list of OIHD, a list of P-OIHD would inform various PSET planning exercises while being more geographically disaggregated. Therefore, the main goal of this report was to ensure that the development of a list of P-OIHD could be done in a consistent way across provinces, standardising the cross-provincial approach.

Moreover, this report also provides a platform for critique (much like other methodological reports before it). With its publication, we believe that critical engagement can hone the approach further and provide a robust way to engage in province-level planning. This continuous engagement is critical to ensure that PSET planning is contextual, timely and, above all, widely accepted, providing an agreed-upon means of moving forward and improving labour market outcomes across the nine South African provinces.

PART 4

Bibliography

Aguiar-Conraria, L., Martins, M., & Soares, M. (2019). Okun's Law Across Time and Frequencies. *Journal of Economic Dynamics and Control* (13).

Ball, L., Furceri, D., Leigh, D., & Loungani, P. (2019). Does One Law Fit All? Cross-Country Evidence on Okun's Law. *Open Economies Review, 30* (5), 841–874.

Capazario, M., & Venter, F. (2020). *The 2020 List of Occupations in High Demand: A Technical Report.* Tshwane: Department of Higher Education and Training.

Department of Corporate Governance and Traditional Affairs (2020, July 3). *District Development Model.* Retrieved from COGTA: https://www.cogta.gov.za/ddm/index.php/about-us/

Gonzalez, N., & Jontxen, A. (2020). *Testing Okun's Law in Developed Countries and the Economic Shocks in Spain.* Universidad del Pais Vasco.

Migration Advisory Council (2020). *Review of the Shortage Occupation List: 2020.* London: Migration Advisory Council.

Migration Advisory Council (2013). *Skilled, Shortage, Sensible: A Full Review of the Recommended Shortage Occupation Lists for the UK and Scotland, a Sunset Clause and the Creative Occupations.* London: Migration Advisory Council.

OECD (2017). *Getting Skills Right: Skills for Jobs Indicators.* Paris: OECD Publishing.

Onakoya, A., & Seyingbo, A. (2020). Economic Growth and Unemployment Nexus: Okun's Two-Version Case for Nigeria, South Africa and the United States of America. *Journal of Economics and Behavioural Studies*, 55–65.

Reddy, V., Rogan, M., Mncwango, B., & Chabane, S. (2018). *Occupations in High Demand in South Africa.* Labour Market Intelligence Partnership.

Sharma, N. (2021, May 10). *Ways to Detect and Remove Outliers.* Retrieved from Towards Data Science: https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba

Department of Higher Education and Training (2015). *Guidelines: Organising Framework for Occupations (OFO).* Tshwane: Department of Higher Education and Training.

Department of Higher Education and Training (2019). *Organising Framework for Occupations.* Tshwane: Department of Higher Education and Training.

Appendix

Appendix 1: Comparison with a national OIHD methodology

The list of P-OIHD has certain methodological similarities to the national list of OIHD, the List of Priority Occupations (LPO) and the Critical Skills List (CSL).

In general, these lists attempt to measure occupational demand and shortage through quantitative analysis and validate that quantitative analysis with stakeholder validation. However, the direct comparator to the provincial list is the national list. As a result, it is important to pinpoint exactly where the national list of OIHD and the list of P-OIHD are similar and where they diverge. This similarity is discussed in Box 2 below.

Box 2 High-level differences between national and provincial lists of OIHD

- The national list of OIHD's quantitative methodology relies on labour market data from the QLFS, the LMDS and Career Junction. Indicators related to wage, employment and vacancy pressure are the key drivers of the analysis.
- These indicators are then combined using Principal Components Analysis (PCA), among other weighting methods, to create a multidimensional index of occupational demand. This approach follows best practice examples[15] and is conducted at the 4-digit/unit group OFO level.
- There are, however, some factors that constrain the use of such an approach at the provincial level:
  1. The number of observations for each Stats SA curated dataset (QLFS and LMDS) at the national level usually hovers between 30 000 and 50 000 individuals. While these datasets are nationally representative, subdividing the data cuts into the number of observations for each province dramatically.[16] Therefore, relying on analysis using the QLFS or LMDS alone would not represent the reality in the provinces given the lack of provincial observations, especially when the analysis would only focus on individuals employed within specific occupational groupings.
  2. The QLFS and LMDS contain the above indicators at the unit group level. However, the data is more reliable at the 3-digit occupational level (especially when cross-tabulating the data).
  3. Finally, the national list of OIHD's methodology is somewhat of a black box for most readers.

For these reasons, the researchers propose the following changes:

- The data analysis should be done at the 3-digit OFO level across all data sources. This approach balances out data reliability and disaggregation more than the 4-digit approach.
- Instead of PCA or other difficult-to-understand weighting schemes, the multi-dimensional index should be estimated by following the approach more recently applied by the Migration Advisory Council. This approach ranks indicators across minor groups according to their relationship with occupational demand. The ranks are then aggregated to get an average rank for each minor group.

15 The methodology for the national list of OIHD used insights from OECD (2017) and the Migration Advisory Council (2013), which combined the indicators using a statistical methodology.

16 For example, Q3 of the QLFS for 2020 sampled 47 167 individuals. Of those, roughly a quarter were sampled from Gauteng and KwaZulu-Natal. This does not match the population distribution in reality.

These and some other differences between the national and provincial methodology are set out in Table 3 below.

TABLE 3: Key differences between national and provincial lists of OIHD

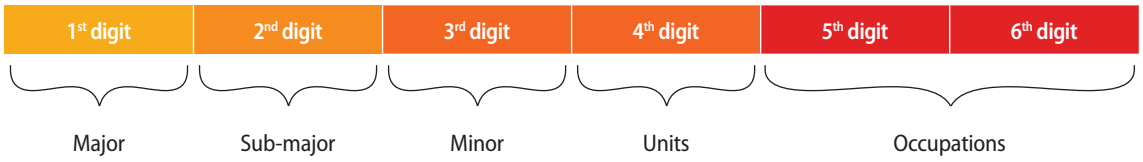
METHODOLOGICAL ASPECT	NATIONAL	PROVINCIAL
Unit of quantitative analysis	4-digit sub-minor group	3-digit minor group
Ranking demand	PCA and equal weighting based on exceeding threshold (median)	Rank each variable and calculate average ranking across variables and dimensions (newest MAC approach)
Treatment of survey weights	Not considered	Analysis based on survey weights, making each list representative
Treatment of outliers	Included in analysis	Excluded from analysis (specifically wage outliers)
Base year	Remained at 2010	Suggesting a rolling 2-year period (most recent data and 2 years prior)
Macroeconomic outlook	Not explicitly considered	Industry-level macroeconomic outlook
Call for evidence	Open-call semi-structured questionnaire	Structured survey with minor groups and relevant occupations matched to pre-identified stakeholders

Appendix 2: Organising Framework for Occupations (OFO)

According to the DHET, the OFO is a classification framework that establishes "a common language for talking about occupations". As the 2018 technical report for the national list OIHD puts it:

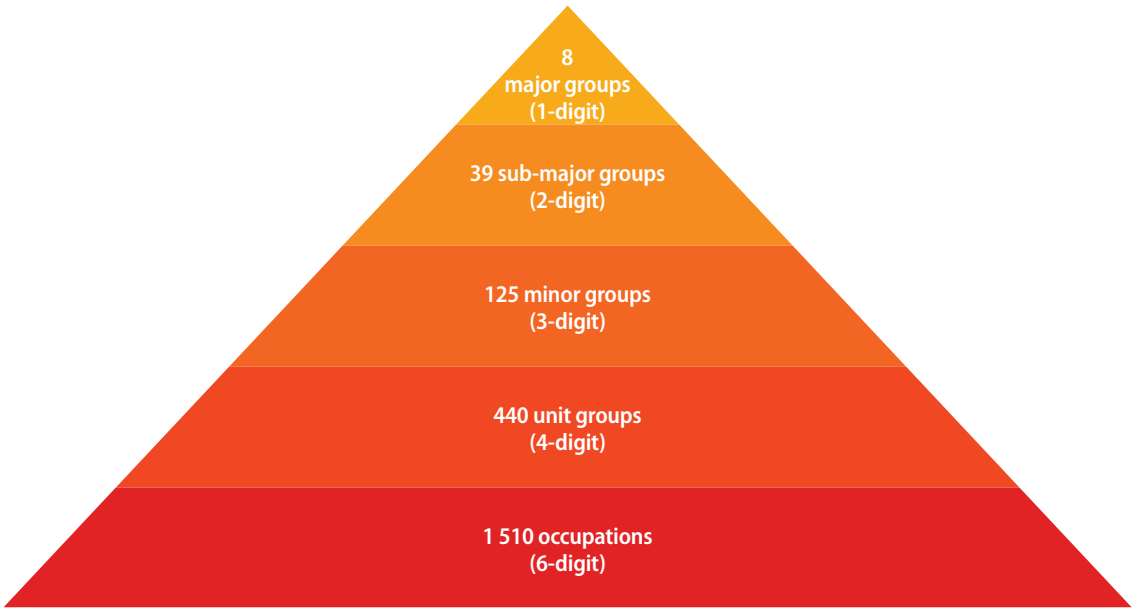
> The OFO is a skill-based, coded classification system of occupations which is used by the DHET for identifying, reporting, and monitoring skills demand and supply in the labour market. It captures almost all occupations in the country and classifies them by skill level and skill specialisation. (Reddy, Rogan, Mncwango & Chabane, 2018, p. 10)

The framework gives each occupation a 6-digit numeric code. Occupations are then grouped into 4-digit unit groups, 3-digit minor groups, 2-digit sub-major groups and, finally, 1-digit major groups. Within the six digits of the occupation, each number refers to a specific grouping.



As shown in Figure 13, the eight 1-digit major groups together encapsulate the entire set of 1 510 6-digit occupations.

FIGURE 13: Number of classifications in the OFO



Source: (The Department of Higher Education and Training, 2019)

The eight major groups indicate the broad area of specialisation of occupations. Figure 14 maps these major groups to the National Qualifications Framework (NQF) levels. As the major groups move from 1 to 8, the NQF levels required to ply a particular occupation decrease.

FIGURE 14: Relationship between OFO major categories and NQF level

<table>
<tr><th>NSDS (level of skill required for a given NQF)</th><th>NQF</th><th colspan="4">OFO MAJOR GROUP</th></tr>
<tr><td rowspan="4">High</td><td>10</td><td colspan="2" rowspan="4">2
Professionals</td><td rowspan="4">1
Managers</td></tr>
<tr><td>9</td></tr>
<tr><td>8</td></tr>
<tr><td>7</td></tr>
<tr><td rowspan="2">Intermediate</td><td>6</td><td colspan="2" rowspan="2">3
Technicians and associate professionals</td></tr>
<tr><td>5</td><td rowspan="3">4
Clerical support workers</td><td rowspan="3">5
Service and sales workers</td><td rowspan="3">6
Skilled agricultural, forestry, fishery, craft and related trade workers</td><td rowspan="3">7
Plant and machine operators and assemblers</td></tr>
<tr><td rowspan="3">Entry</td><td>4</td></tr>
<tr><td>3</td></tr>
<tr><td>2</td><td colspan="4" rowspan="2">8
Elementary occupations</td></tr>
<tr><td>1</td></tr>
</table>

Source: Department of Higher Education and Training (2015)

The OFO is foundational to the methodology outlined below. Most importantly, although labour demand should ideally be analysed at the occupational level (digit level 6 of the OFO), the data required to do this does not exist. Therefore, the methodology starts with an analysis of labour market and macroeconomic prospects for each 3-digit minor group. This evidence, along with a value chain analysis conducted by the provinces (as will be discussed further on), will provide both quantitative and qualitative information regarding the degree to which 3-digit minor groups show signs of demand.

Once all 3-digit minor groups are correctly identified, all matched 6-digit occupations will be compiled into a list for validation. Stakeholders in the provincial skills landscape, including industry-specific representatives, will then be surveyed using Likert scale (rating-type) questions to identify which of these 6-digit occupations within the minor groups are the most in demand in each province.

Appendix 3: Step-by-step guide to employing an ARIMA model for forecasting

- **Step 1: Import data into EViews**
  - Click File ➡ New ➡ Workfile
  - Select the frequency of data you're using and type in the beginning date of the dataset. In the end date area, extend the period by as far as you need for the forecasts. Click OK.
  - Click Quick ➡ Empty Group (edit series) and copy-paste your data from Excel, making sure to include the heading in lowercase without any spaces between heading words. Escape the sheet once you've pasted the data, and your variable will appear in your Workfile window.
- **Step 2: Estimate a correlogram**
  - Double click on your variable of interest.
  - Click View ➡ Correlogram … ➡ Level and press OK
  - You will now have two graphs in front of you, an autocorrelation and partial correlation function (ACF and PACF), like the one shown below.

FIGURE 15: Autocorrelation and partial autocorrelation function example

Date: 02/17/22 Time: 10:51am
Sample: 1998 2023
Included observations: 22

AUTOCORRELATION	PARTIAL CORRELATION		AC	PAC	Q-STAT	PROB
		1	0.790	0.790	15.691	0.000
		2	0.551	−0.194	23.713	0.000
		3	0.315	−0.143	26.470	0.000
		4	0.108	−0.097	26.814	0.000
		5	−0.094	−0.183	27.086	0.000
		6	−0.284	−0.192	29.746	0.000
		7	−0.363	0.068	34.396	0.000
		8	−0.396	−0.083	40.296	0.000
		9	−0.328	0.118	44.674	0.000
		10	−0.320	−0.266	49.173	0.000
		11	−0.333	−0.193	54.496	0.000
		12	−0.276	0.073	58.511	0.000

- **Step 3: Interpret the correlogram and estimate the corresponding equation**
  - Interpreting the ACF and PACF simply requires two things:
    - First, that you count the number of spikes in each graph that extend above the dotted line.
    - Second, that you assess whether the ACF or PACF "geometrically decay". A geometric decay simply means that the bars in the graph look like they're cycling like a sine wave or decreasing towards zero, without any strange spikes.
  - PACFs relate to how many AR components our forecast regression will have.
    - In the PACF in the example, one big spike occurs above the dotted line, and the PACG does not geometrically decay.
    - This signals that our forecast equation will have only one AR component.
  - ACFs relate to how many MA components our forecast regression will have.
    - In the ACF in the example, there are two spikes above the dotted line, but the graphic looks like its cycling like a sine wave. If there were two spikes without that geometric decay, we would include two MA components in the regression. However, because of this geometric decay, there is no need to include MA components in the forecast equation.
  - The corresponding GVA forecast equation will then simply be:

$$GVA_{i,t} = \beta_1 + \beta_2 \, GVA_{i,t-1}$$

- **Step 4: Specify this equation in your statistical package and produce forecasts**
  - To estimate the first equation in EViews and produce forecasts, follow the next steps:
    - Click Quick ➡ Estimate Equation…
    - Type in the regression specification of your choice based on the Box-Jenkins methodology. In the case before, one would write the following in the command window.

```
gva c gva(-1)
```

    - Click Forecast and change the forecast sample to the out-of-sample dates that you'd like to forecast into.
    - A forecast graph, along with a new variable, will appear in your Workfile space. The new variable (gvaf) is the forecast for the GVA value for the specified time-period, stored in an Excel format.
- **Step 5: Repeat this process for all industries' GVA**

Notes

Notes

Design and Layout: Ink Design Publishing Solutions, Cape Town, www.inkdesign.co.za

DPRU CONTACTS

Programme Leader: Prof. Haroon Bhorat – haroon.bhorat@uct.ac.za
Programme Manager: Ms Kezia Lilenstein – kezia.lilenstein@uct.ac.za

DHET CONTACTS

Programme Leader: Ms M. Khuluvhe – Khuluvhe.M@dhet.gov.za
Project Secretariat: Ms M. Ramasodi – Ramasodi.M@dhet.gov.za