

Enquête sur les exigences linguistiques des entreprises et des municipalités du Québec en 2023

Rapport méthodologique

Pour tout renseignement concernant l'Institut de la statistique du Québec (ISQ) et les données statistiques dont il dispose, s'adresser à :

Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4

Téléphone :
418 691-2401
1 800 463-4090 (Canada et États-Unis)

Site Web : statistique.quebec.ca

Ce document est disponible seulement en version électronique.

Dépôt légal
Bibliothèque et Archives nationales du Québec
1^{er} trimestre 2025
ISBN 978-2-555-00303-3 (en ligne)

© Gouvernement du Québec, Institut de la statistique du Québec, 2025

Toute reproduction autre qu'à des fins de consultation personnelle est interdite sans l'autorisation du gouvernement du Québec.
statistique.quebec.ca/fr/institut/nous-joindre/droits-auteur-permission-reproduction

Janvier 2025

Rapport réalisé à l'Institut de la statistique
du Québec par : Layla Linteau, statisticienne

Sous la coordination de : Luc Belleau, coordonnateur

Sous la direction de : Éric Gagnon, directeur

Révision linguistique: Direction de la diffusion et des communications

Travaux méthodologiques : Mathieu Baillargeon, statisticien
Layla Linteau, statisticienne

Enquête sous la responsabilité de : Direction de la méthodologie

Enquête financée par : Office québécois de la langue française

Pour tout renseignement concernant
le contenu de ce document : Direction de la méthodologie
Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4

Téléphone :
418 691-2410 ou
1 800 463-4090
(sans frais d'appel au Canada et aux États-Unis)
Télécopieur : 418 643-4129

Site Web : statistique.quebec.ca

Notice bibliographique suggérée

LAYLA, Linteau (2025). *Enquête sur les exigences linguistiques des entreprises et des municipalités du Québec en 2023 – Rapport méthodologique*, [En ligne], Québec, Institut de la statistique du Québec, 28 p. [statistique.quebec.ca/fr/fichier/enquete-exigences-linguistiques-entreprises-municipalites-2023-metho.pdf].

Table des matières

Introduction	6
La méthodologie	7
1 Le plan d'échantillonnage	8
1.1 Population visée et base de sondage	8
1.1.1 Volet auprès des établissements	8
1.1.2 Volet auprès des municipalités et des arrondissements de la Ville de Montréal	8
1.2 Méthode de sélection, taille et répartition de l'échantillon	9
2 La collecte de données	11
2.1 Instruments de collecte	11
2.2 Mode de collecte	11
2.3 Déroulement de la collecte	12
2.4 Taux de réponse	12
2.4.1 Volet auprès des municipalités et des arrondissements de la Ville de Montréal	12
2.4.2 Volet auprès des établissements	12
3 La pondération	15
3.1 Volet auprès des municipalités et des arrondissements de la Ville de Montréal	15
3.2 Volet auprès des établissements	16
4 Les méthodes d'estimation et d'analyse des résultats	17
4.1 Types d'estimation	17
4.2 Tests d'association	17

5	L'évaluation de la qualité statistique des résultats	18
5.1	Erreurs dues à l'échantillonnage	18
5.1.1	Intervalle de confiance	18
5.1.2	Coefficient de variation	19
5.2	Erreurs non dues à l'échantillonnage	20
5.2.1	Taux de réponse global et taux de non-réponse partielle	20
5.2.2	Réponses aux questions	20
5.2.3	Couverture de la base de sondage du volet auprès des établissements	20
6	La confidentialité	22
7	L'appréciation globale	23
Annexe 1	Population de l'enquête	24
Annexe 2	Résultats de collecte	27

Introduction

L'Institut de la statistique du Québec (Institut) a été mandaté par l'Office québécois de la langue française (Office) pour réaliser en 2023 la deuxième édition¹ de *l'Enquête sur les exigences linguistiques des entreprises et des municipalités du Québec*. L'objectif principal de cette enquête consiste à mieux connaître quelles sont les connaissances linguistiques exigées ou souhaitées chez les employés et

employées. Les résultats de l'enquête permettront de mieux connaître les habitudes des employeurs québécois en matière d'exigences linguistiques pour la sélection d'un candidat ou d'une candidate lors d'un processus d'embauche, de mutation ou de promotion. Ce rapport présente la méthodologie mise en œuvre ainsi qu'une évaluation de la qualité de cette enquête réalisée par l'Institut.

1. Il est à noter que la première édition de l'enquête portait un titre légèrement différent, soit *Enquête sur les exigences linguistiques auprès des entreprises, des municipalités et des arrondissements de Montréal*.

La méthodologie

La méthodologie de l'*Enquête sur les exigences linguistiques des entreprises et des municipalités du Québec en 2023* est présentée dans les sections qui suivent. On y traite du plan d'échantillonnage, des instruments et du mode de collecte, de la collecte des données, de l'estimation, de l'évaluation statistique de l'enquête et de la confidentialité. On y trouve également une section faisant état de l'appréciation globale de l'enquête.

1 Le plan d'échantillonnage

1.1 Population visée et base de sondage

La population visée par l'enquête est constituée de deux types d'employeurs québécois, soit les entreprises² d'une part, soit les municipalités et les arrondissements de la Ville de Montréal d'autre part.

1.1.1 Volet auprès des établissements

La population visée par ce volet de l'enquête est constituée de tous les établissements québécois employant cinq personnes et plus des secteurs définis par les codes du Système de classification des industries de l'Amérique du Nord (SCIAN), c'est-à-dire ceux de la construction, de la fabrication et des services du secteur privé (à l'exception des soins de santé et de l'assistance sociale). Les établissements situés hors Québec sont exclus de la base de sondage.

L'Institut a constitué la base de sondage à partir de la Banque de données du registre des entreprises (BDRE) de Statistique Canada. La BDRE présente de nombreux avantages par rapport à d'autres sources d'information sur les entreprises. Les plus importants sont sans doute sa mise à jour régulière et le fait qu'elle soit constituée

spécialement pour la réalisation d'enquêtes. Les autres avantages sont la désagrégation de la structure des entreprises jusqu'à l'emplacement, de même que la présence d'une classification industrielle à jour et assez précise, le SCIAN. On y trouve aussi une variable catégorielle estimant le nombre d'employés et employées de chaque établissement, qui peut s'avérer précieuse pour l'élaboration du plan de sondage. La base de sondage contient 104 090 établissements et la liste de leurs coordonnées.

1.1.2 Volet auprès des municipalités et des arrondissements de la Ville de Montréal

La population visée pour ce second volet est constituée de toutes les municipalités de 5 000 habitants et habitantes ou plus.

Pour ce volet, plusieurs sources de données ont été nécessaires pour la création de la base de sondage. En effet, l'Institut a utilisé deux fichiers du site Internet Données Québec afin de dresser une liste de 189 municipalités et des 19 arrondissements de la Ville de Montréal. Ces travaux ont permis également de colliger les coordonnées des personnes en mesure de répondre aux questions de l'enquête.

2. Il est important de mentionner que cette enquête est faite auprès des établissements, car c'est là que les personnes-ressources sont le plus susceptibles de connaître les exigences linguistiques lors de l'embauche.

1.2 Méthode de sélection, taille et répartition de l'échantillon

L'Office désirait obtenir des renseignements fiables selon le regroupement de secteurs d'activité, la taille de l'établissement et le regroupement de régions. Les unités qui composent le volet auprès des municipalités et des arrondissements ont été recensées avec un taux de réponse attendu de 90 %, soit pour un total de 187 municipalités et arrondissements répondants visés. Pour le volet auprès des établissements, un plan de sondage stratifié a été proposé et les strates ont été définies comme étant le croisement :

- du secteur d'activité (6 regroupements)³ ;
- de la taille de l'établissement (4 tailles) ;
- de la région (2 catégories).

Au total, 48 strates d'échantillonnage ont été construites. À l'intérieur de chacune, un échantillon probabiliste a été tiré de façon aléatoire. La taille de l'échantillon a été fixée à 3 420 établissements, avec un taux d'admissibilité estimé de 85 % et un taux de réponse visé de 80 %, pour un objectif de 2 303 établissements répondants attendus. Le tableau 1 présente les données relatives à la taille de la population, à la taille de l'échantillon et au nombre d'établissements répondants attendu, selon les variables de stratification.

Tableau 1

Nombre d'établissements dans la population, nombre d'établissements dans l'échantillon et nombre d'établissements répondants attendu, selon les variables de stratification au volet auprès des établissements

Secteur	Nombre d'employé(e)s	Région	Taille de la population	Taille de l'échantillon	Nombre d'établissements répondants attendu
Fabrication et services à forte concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	2 334	83	53
		Autres régions	4 114	133	85
	20 à 29 employé(e)s	Île de Montréal	383	35	22
		Autres régions	554	40	26
	30 à 49 employé(e)s	Île de Montréal	355	44	28
		Autres régions	448	55	35
	50 employé(e)s ou plus	Île de Montréal	596	79	51
		Autres régions	588	65	42
Fabrication à concentration moyenne de connaissances	5 à 19 employé(e)s	Île de Montréal	852	70	48
		Autres régions	3 130	98	67
	20 à 29 employé(e)s	Île de Montréal	195	30	20
		Autres régions	721	48	33
	30 à 49 employé(e)s	Île de Montréal	209	34	23
		Autres régions	756	73	50
	50 employé(e)s ou plus	Île de Montréal	367	40	27
		Autres régions	1 293	96	65

Suite à la page 10

3. Voir l'annexe 1 pour les secteurs inclus dans les regroupements.

Tableau 1 (suite)

Nombre d'établissements dans la population, nombre d'établissements dans l'échantillon et nombre d'établissements répondants attendu, selon les variables de stratification au volet auprès des établissements

Secteur	Nombre d'employé(e)s	Région	Taille de la population	Taille de l'échantillon	Nombre d'établissements répondants attendu
Services à concentration moyenne de connaissances	5 à 19 employé(e)s	Île de Montréal	4 663	148	101
		Autres régions	10 781	123	84
	20 à 29 employé(e)s	Île de Montréal	870	27	18
		Autres régions	1 789	43	29
	30 à 49 employé(e)s	Île de Montréal	796	40	27
		Autres régions	1 380	61	41
	50 employé(e)s ou plus	Île de Montréal	1 487	67	46
		Autres régions	1 858	67	46
Fabrication à faible concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	270	81	55
		Autres régions	4 098	206	140
	20 à 29 employé(e)s	Île de Montréal	38	18	12
		Autres régions	437	52	35
	30 à 49 employé(e)s	Île de Montréal	31	18	12
		Autres régions	353	73	50
	50 employé(e)s ou plus	Île de Montréal	56	24	16
		Autres régions	532	103	70
Services à faible concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	7 093	206	140
		Autres régions	23 406	155	105
	20 à 29 employé(e)s	Île de Montréal	1 322	24	16
		Autres régions	4 110	50	34
	30 à 49 employé(e)s	Île de Montréal	935	27	18
		Autres régions	3 340	74	50
	50 employé(e)s ou plus	Île de Montréal	1 134	39	27
		Autres régions	3 271	93	63
Construction	5 à 19 employé(e)s	Île de Montréal	1 063	89	61
		Autres régions	9 199	206	140
	20 à 29 employé(e)s	Île de Montréal	133	18	12
		Autres régions	940	53	36
	30 à 49 employé(e)s	Île de Montréal	121	20	14
		Autres régions	738	75	51
	50 employé(e)s ou plus	Île de Montréal	152	24	16
		Autres régions	799	93	63
Total pour le volet auprès des établissements			104 090	3 420	2 303

2 La collecte de données

2.1 Instruments de collecte

Le questionnaire de la présente enquête a été élaboré par l'Office en collaboration avec l'Institut. La programmation des deux versions (téléphonique et électronique) du questionnaire a été réalisée par l'Institut.

Le questionnaire de l'enquête comprend 32 questions réparties dans 5 sections :

1. Structure de l'entreprise (uniquement pour le volet auprès des établissements) ;
2. Caractéristiques de la dernière embauche, promotion ou mutation ;
3. Exigences linguistiques ;
4. Évaluation des compétences linguistiques et sélection de la candidature ;
5. Processus d'embauche de l'organisation.

Le questionnaire a subi quelques changements pour la deuxième édition de l'enquête. Il y a notamment cinq questions qui ont été ajoutées dans la section Exigences linguistiques afin de distinguer les compétences exigées de celles souhaitées.

2.2 Mode de collecte

La collecte de l'*Enquête sur les exigences linguistiques des entreprises et des municipalités du Québec en 2023* a consisté en une collecte multimode, c'est-à-dire que la collecte de données a été réalisée par l'intermédiaire de questionnaires Web et de relances téléphoniques. Soulignons que cette enquête était à déclaration obligatoire.

Notons que le questionnaire pouvait être rempli par téléphone si la personne répondante préférait ce mode de collecte. Un filtre téléphonique préalable à la collecte a validé l'admissibilité des établissements et a permis d'identifier la meilleure personne pour répondre au questionnaire. Cette étape préalable permet d'optimiser l'ensemble des activités de collecte et d'assurer la qualité des données recueillies.

Dans un souci d'uniformité entre les modes de collecte, les questionnaires étaient identiques afin d'éviter de créer un effet de mode dans les données.

Une lettre d'invitation à participer à l'enquête a été envoyée aux employeurs échantillonnés. Dans celle-ci, une description de l'enquête, un identifiant unique ainsi que l'adresse Internet pour accéder au questionnaire électronique étaient communiqués. Une semaine plus tard, un premier courriel de relance a été envoyé à ceux qui n'avaient pas, à ce jour, répondu au questionnaire électronique. La collecte téléphonique a ensuite été effectuée auprès des personnes qui n'avaient pas rempli le questionnaire électronique ou qui préféraient répondre par téléphone.

La collecte électronique a été effectuée au moyen d'un système d'interview sur le Web assisté par ordinateur (IWAQ), tandis que la collecte téléphonique a été réalisée au moyen d'un système d'interview téléphonique assisté par ordinateur (ITAO).

2.3 Déroutement de la collecte

La collecte auprès des employeurs échantillonnés a été effectuée du 5 octobre 2023 au 8 janvier 2024. Les questions du filtre, préalable à la collecte, ont permis de caractériser l'établissement sélectionné et d'en déterminer l'admissibilité. Le filtre a permis également d'identifier la personne qui pouvait répondre au questionnaire et d'obtenir ses coordonnées. Par la suite, les documents ont été expédiés par courriel (ou par la poste) aux personnes identifiées. Cet envoi comprenait une lettre d'invitation, un questionnaire, un lien et un identifiant pour accéder au questionnaire Web.

Rappelons que la collecte téléphonique était faite auprès des personnes qui n'avaient pas rempli le questionnaire électronique ou qui préféraient répondre au questionnaire téléphonique. Des courriels de relance ont été envoyés entre le 18 octobre et le 6 décembre 2023 aux employeurs n'ayant pas rempli le questionnaire électronique. Des relances téléphoniques ont également été faites à quelques reprises pendant la collecte. Le 20 novembre, les documents ont été envoyés par la poste aux employeurs qui n'avaient pas été joints. La période de collecte téléphonique était du lundi au vendredi, entre 8 h 30 et 16 h 30.

Une équipe de sept intervieweurs et intervieweuses a reçu une formation le 2 octobre 2023 pour procéder à la collecte téléphonique. Cette formation portait sur les objectifs de l'enquête, le contenu du questionnaire, la procédure de validation des coordonnées des employeurs et les codes de résultat. La formation comportait une partie théorique, des mises en situation, la présentation du questionnaire électronique et une période de mise en pratique pour se familiariser avec les outils de collecte ainsi que la documentation disponible.

La durée moyenne pour remplir le questionnaire électronique était de 10 minutes. Le temps total moyen pour mener un dossier à terme, y compris les tentatives de connexion et le temps requis pour remplir le questionnaire électronique, était de 28 minutes. Notons que la personne sélectionnée devait entrer son identifiant unique reçu dans la lettre d'invitation, créer un mot de passe et valider son identité pour accéder au

questionnaire électronique. De plus, celui-ci pouvait être rempli en français ou en anglais, au choix de la personne répondante.

Pour la collecte téléphonique, sept appels en moyenne, à des heures et à des jours différents, étaient effectués pour joindre la personne sélectionnée et mener l'entrevue à bien. Les entrevues étaient faites en français ou en anglais, au choix de la personne répondante. La durée moyenne pour une entrevue (le temps de remplir le questionnaire téléphonique) était de 22 minutes. Le temps total moyen consacré aux dossiers menés à terme, y compris l'entrevue et les tentatives de contact, était de 61 minutes.

2.4 Taux de réponse

2.4.1 Volet auprès des municipalités et des arrondissements de la Ville de Montréal

Le taux de réponse final pour ce volet est de 94,7 %, soit de près de 15 points de pourcentage supérieur au taux attendu. Environ 98 % des questionnaires ont été remplis par Internet, tandis qu'environ 2 % des questionnaires ont été remplis lors de la collecte téléphonique.

2.4.2 Volet auprès des établissements

Au total, 2 483 établissements ont rempli le questionnaire, alors que le nombre attendu était de 2 303 (tableau 2). Environ 94,5 % des questionnaires ont été remplis par Internet, environ 4,5 % lors de la collecte téléphonique et moins de 1 % après la réception du questionnaire par la poste ou par télécopieur.

Comme les établissements n'ont pas tous la même probabilité d'être choisis, il est intéressant de considérer un taux de réponse établi sur la base des données pondérées⁴ par l'inverse de cette probabilité. Ce taux de réponse pondéré final pour le volet auprès des établissements a été de 85,3 % (tableau 2), ce qui est plus élevé que celui attendu, fixé à 80 %. Le taux d'admissibilité pondéré a été de 83,3 %.

4. La pondération sera abordée à la section 3.

On remarque au tableau 2 que le nombre d'établissements répondants attendu a été atteint pour la majorité des croisements des trois variables de stratification. Toutefois, on peut remarquer que le plus faible taux de réponse pondéré est pour la combinaison « Fabrication à faible concentration de connaissances » × « 50 employé(e)s ou plus » × « Île de Montréal » et la combinaison « Services à faible concentration de connaissances » × « 20 à 29 employé(e)s » × « Île de Montréal ». Mentionnons finalement que les établissements de la région de Montréal ont été proportionnellement moins nombreux à répondre que les établissements situés dans les autres régions (taux de réponse pondérés de 80,4 % et de 86,2 %, respectivement), ce qui est cohérent avec certains résultats observés dans d'autres enquêtes auprès des entreprises.

Même si le nombre d'établissements répondants est plus élevé que celui attendu pour l'ensemble du Québec et pour la quasi-totalité des trois variables de stratification, il se peut que la non-réponse obtenue engendre un biais dans les estimations. Ce biais a été amoindri par la pondération.

Tableau 2

Nombre d'établissements répondants, nombre d'établissements répondants attendu, taux de réponse et taux d'admissibilité pondérés, selon les trois variables de stratification

Secteur	Nombre d'employé(e)s	Région	Nombre d'établissements répondants	Nombre d'établissements répondants attendu	Taux de réponse pondéré	Taux d'admissibilité pondéré
			n		%	
Fabrication et services à forte concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	47	53	80,7	75,9
		Autres régions	92	85	89,5	79,7
	20 à 29 employé(e)s	Île de Montréal	25	22	74,3	97,1
		Autres régions	37	26	95,0	97,5
	30 à 49 employé(e)s	Île de Montréal	33	28	75,0	100
		Autres régions	48	35	90,9	96,4
	50 employé(e)s ou plus	Île de Montréal	62	51	82,3	96,2
		Autres régions	56	42	90,8	95,4
Fabrication à concentration moyenne de connaissances	5 à 19 employé(e)s	Île de Montréal	36	48	75,7	75,7
		Autres régions	62	67	88,8	74,5
	20 à 29 employé(e)s	Île de Montréal	28	20	93,3	100
		Autres régions	36	33	81,3	93,8
	30 à 49 employé(e)s	Île de Montréal	27	23	79,4	100,0
		Autres régions	61	50	90,4	93,2
	50 employé(e)s ou plus	Île de Montréal	33	27	85,0	97,5
		Autres régions	86	65	91,7	97,9

Suite à la page 14

Tableau 2 (suite)

Nombre d'établissements répondants, nombre d'établissements répondants attendu, taux de réponse et taux d'admissibilité pondérés, selon les trois variables de stratification

Secteur	Nombre d'employé(e)s	Région	Nombre d'établissements répondants	Nombre d'établissements répondants attendu	Taux de réponse pondéré	Taux d'admissibilité pondéré
			n		%	
Services à concentration moyenne de connaissances	5 à 19 employé(e)s	Île de Montréal	84	101	79,7	77,0
		Autres régions	72	84	87,0	71,5
	20 à 29 employé(e)s	Île de Montréal	20	18	77,8	96,3
		Autres régions	33	29	86,0	90,7
	30 à 49 employé(e)s	Île de Montréal	33	27	90,0	92,5
		Autres régions	46	41	85,3	90,2
	50 employé(e)s ou plus	Île de Montréal	50	46	79,1	95,5
		Autres régions	55	46	89,6	92,5
Fabrication à faible concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	54	55	81,5	85,2
		Autres régions	117	140	85,0	71,8
	20 à 29 employé(e)s	Île de Montréal	13	12	83,3	88,9
		Autres régions	43	35	84,6	98,1
	30 à 49 employé(e)s	Île de Montréal	15	12	88,9	94,4
		Autres régions	64	50	93,2	94,5
	50 employé(e)s ou plus	Île de Montréal	17	16	70,8	100,0
		Autres régions	92	70	93,2	96,1
Services à faible concentration de connaissances	5 à 19 employé(e)s	Île de Montréal	122	140	80,1	79,1
		Autres régions	103	105	85,2	81,2
	20 à 29 employé(e)s	Île de Montréal	17	16	70,8	100,0
		Autres régions	39	34	84,0	94,0
	30 à 49 employé(e)s	Île de Montréal	22	18	85,2	96,3
		Autres régions	58	50	83,8	94,6
	50 employé(e)s ou plus	Île de Montréal	28	27	82,1	89,7
		Autres régions	80	63	89,3	96,7
Construction	5 à 19 employé(e)s	Île de Montréal	56	61	85,4	77,8
		Autres régions	139	140	87,7	79,6
	20 à 29 employé(e)s	Île de Montréal	14	12	94,4	83,3
		Autres régions	44	36	86,8	96,2
	30 à 49 employé(e)s	Île de Montréal	16	14	100,0	80,0
		Autres régions	64	51	86,7	98,7
	50 employé(e)s ou plus	Île de Montréal	18	16	91,7	83,3
		Autres régions	86	63	96,8	95,7
Total pour le volet auprès des établissements			2 483	2 303	85,3	83,3

3 La pondération

La pondération a pour objectif fondamental de permettre l'inférence des résultats observés à la population étudiée. En fait, le poids accordé à un employeur reflète le nombre d'établissements que celui-ci « représente » dans la population. Ce poids est donc fonction de deux réalités. Initialement, il constitue l'inverse de la probabilité de sélection d'un établissement dans sa strate. Ensuite, il tient compte de la structure de l'admissibilité des établissements ainsi que de la réponse obtenue dans la classe de pondération à laquelle l'établissement appartient, en supposant que le profil des établissements non répondants de cette classe de pondération est semblable à celui des établissements répondants. Cette dernière étape permet de constituer le poids final.

3.1 Volet auprès des municipalités et des arrondissements de la Ville de Montréal

Rappelons que les unités constituant ce volet ont été recensées. Ainsi, leur probabilité de sélection est de 100 % et leur poids est de 1 à cette étape. Le taux d'admissibilité étant de 100 %, aucun ajustement pour l'admissibilité n'a été effectué. Un ajustement pour la non-réponse totale a été nécessaire. On a obtenu 197 municipalités et arrondissements répondants et 11 non-répondants. Les classes de pondération sont composées selon la taille de la municipalité. Les arrondissements de la Ville de Montréal ont tous été mis dans la même classe de pondération, peu importe leur taille.

Tableau 3

Taux de réponse et taux d'admissibilité pondérés, selon la stratification

Classe de pondération	Taux d'admissibilité pondéré	Taux de réponse pondéré
	%	
Arrondissements de Montréal	100,0	84,2
Municipalités de 5 000 à 9 999 habitant(e)s	100,0	92,3
Municipalités de 10 000 à 29 999 habitant(e)s	100,0	98,5
Municipalités de 30 000 habitant(e)s ou plus	100,0	97,8

3.2 Volet auprès des établissements

Pour effectuer une pondération adéquate, il faut être en mesure d'estimer le nombre total d'établissements inadmissibles au sein de la population. Dans la présente enquête, on a obtenu 2 483 établissements répondants admissibles, 450 établissements inadmissibles et 487 établissements non répondants. Seuls ces derniers possèdent un statut d'admissibilité inconnu. Pour cette raison, l'hypothèse suivante a été émise : parmi ces 487 non-répondants à l'enquête, certains peuvent être considérés comme admissibles, d'autres comme inadmissibles. On peut déterminer un nombre de non-répondants inadmissibles en examinant la nature de la non-réponse par strate. Une fois déterminé, ce nombre additionné de celui des établissements répondants inadmissibles constitue le nombre total d'établissements inadmissibles de l'échantillon. Pour ce faire, on suppose que le profil des établissements non répondants est semblable à celui des répondants par classe de pondération. Une classe de pondération est un regroupement d'établissements ayant des caractéristiques similaires qui ont une influence sur le statut de réponse (répondant ou non-répondant) des établissements. Dans ce cas, les classes de pondération sont les mêmes que les strates d'échantillonnage.

4 Les méthodes d'estimation et d'analyse des résultats

Les estimations produites par l'Institut dans le cadre de la présente enquête respectent les particularités du plan de sondage et sont exposées sous forme de tableaux de fréquences et de tableaux croisés. Pour le volet auprès des établissements, les estimations sont produites pour le Québec, selon le regroupement de secteurs d'activité, la catégorie de taille des établissements et l'emplacement de ceux-ci (île de Montréal ou hors de l'île de Montréal). Pour le volet auprès des municipalités et des arrondissements, les estimations sont produites pour le Québec, selon la catégorie du nombre d'habitants et d'habitantes et l'emplacement de l'unité (île de Montréal ou hors de l'île de Montréal).

4.1 Types d'estimation

Les estimations sont de deux types : l'estimation des paramètres (proportion ou moyenne) et l'estimation de la précision associée à chacun des paramètres. Elles sont produites avec les données pondérées, de sorte qu'elles puissent être inférées à la population visée. Les marges d'erreur, l'intervalle de confiance⁵ et le coefficient de variation associé (voir la section 5.1.2) sont calculés selon le plan de sondage, à l'aide des logiciels SAS et SUDAAN.

4.2 Tests d'association

Un test global du khi-deux⁶ a été effectué pour chacun des tableaux croisés produits par l'Institut et remis à l'Office, selon le plan d'analyse. Ce test permet de voir si une association existe entre la variable analysée et la variable de croisement. Le seuil de signification statistique a été fixé à 5 % pour tous les tests.

Les analyses sont uniquement bivariées, c'est-à-dire que chaque question est mise en relation avec une variable de croisement à la fois. Sauf exception, chaque question est analysée selon le secteur d'activité regroupé, la taille de l'établissement et l'emplacement de celui-ci (île de Montréal ou hors de l'île de Montréal).

Dans le cas où le test du khi-deux était significatif, c'est-à-dire lorsqu'une association entre une question et une variable de croisement était détectée, on examinait s'il y avait un chevauchement dans les intervalles de confiance entre les différentes modalités de la variable de croisement. Au besoin, des tests supplémentaires de différence de proportions deux à deux étaient effectués afin de pouvoir situer précisément dans quelles cellules du tableau se manifestaient ces différences. Ces tests sont basés sur une statistique de Wald calculée à partir de la différence des transformations *logit*⁷ des proportions estimées.

5. Consultez la section 5.1.1 pour plus de détails sur la méthode retenue pour construire les intervalles de confiance.

6. Plus précisément, c'est la statistique F (version log-linéaire) correspondant à un ajustement de Satterthwaite du test du khi-deux qui est utilisée.

7. La transformation « logit » = $\log(p/(1-p))$ est utilisée comme transformation normalisatrice dans la construction d'intervalles de confiance afin d'améliorer leurs taux de couverture, en particulier pour de petites proportions (Korn et Graubard, *Analysis of Health Surveys*, New York, John Wiley & Sons, 1999).

5 L'évaluation de la qualité statistique des résultats

Tout résultat tiré d'enquêtes par échantillon ou par recensement est potentiellement entaché d'erreurs dues ou non à l'échantillonnage. Les sections suivantes décrivent les efforts qui ont été investis afin de limiter ces deux types de sources d'erreurs, de les détecter et de les quantifier.

5.1 Erreurs dues à l'échantillonnage

Les erreurs dues à l'échantillonnage sont attribuables au fait que seule une partie de la population visée est échantillonnée aléatoirement. Ces erreurs sont contrôlables *a priori* à l'étape de la définition du plan de sondage et de la détermination de la taille de l'échantillon. De plus, elles sont quantifiables *a posteriori* et permettent de prendre en considération la variabilité des estimations produites lors de leur analyse.

Les intervalles de confiance et le coefficient de variation sont des mesures utilisées afin de juger de la précision d'une estimation.

5.1.1 Intervalle de confiance

L'étendue de l'intervalle de confiance (IC) est l'une des mesures de la précision d'une estimation. Plus l'intervalle est court, plus la valeur du paramètre est circonscrite. L'intervalle de confiance associé à une proportion représente la zone d'incertitude liée à l'estimation. Cette zone d'incertitude est étroitement liée au niveau de confiance choisi. Ainsi, pour une proportion, un intervalle de confiance à un niveau de confiance de 95 % signifie que si on répétait l'enquête 100 fois et qu'à chaque reprise on estimait la proportion et calculait son intervalle de confiance, 95 des 100 intervalles ainsi créés contiendraient la vraie valeur de la proportion dans la population.

Il existe plusieurs types d'intervalles de confiance. Ceux-ci se distinguent par le choix de la loi servant à l'approximation de la distribution de l'estimateur. Les méthodes standards utilisent habituellement la loi normale ou la loi binomiale. Il existe également des méthodes qui permettent de s'affranchir de l'hypothèse selon laquelle les données suivent une loi connue. Pour la présente enquête, les intervalles de confiance produits sont basés sur la méthode *logit*-Wald, qui utilise une transformation *logit* de la proportion et génère un intervalle asymétrique. Ces intervalles sont cohérents avec le test d'égalité de deux proportions, qui utilise également une transformation *logit*. Cet intervalle de niveau $(1 - \alpha)$ est obtenu de la façon suivante :

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) \pm z_{1-\alpha/2} \sqrt{\hat{v}\left(\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)\right)}$$

En utilisant la méthode du Delta, on peut démontrer que :

$$\hat{v}\left(\ln\left(\frac{\hat{p}}{1-\hat{p}}\right)\right) = \left(\frac{1}{\hat{p}(1-\hat{p})}\right)^2 \hat{v}(\hat{p})$$

L'intervalle de confiance de la fonction *logit* (\hat{p}) est donc délimité par les bornes inférieure et supérieure suivantes :

$$I_{logit} = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right) - z_{1-\alpha/2} \left(\frac{1}{\hat{p}(1-\hat{p})}\right) \sqrt{\hat{v}(\hat{p})}$$

$$S_{logit} = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right) + z_{1-\alpha/2} \left(\frac{1}{\hat{p}(1-\hat{p})}\right) \sqrt{\hat{v}(\hat{p})}$$

Par transformation inverse, les bornes de l'intervalle de confiance pour \hat{p} sont définies comme suit :

$$I_{\hat{p}} = \frac{\exp(I_{logit})}{1 + \exp(I_{logit})}$$

$$S_{\hat{p}} = \frac{\exp(S_{logit})}{1 + \exp(S_{logit})}$$

Le niveau de confiance utilisé pour les intervalles produits à partir des données de cette enquête est de 95 %.

5.1.2 Coefficient de variation

Le coefficient de variation (CV) est une mesure qui permet de quantifier la précision relative d'une estimation et de quantifier l'erreur d'échantillonnage. Il se définit comme suit :

$$CV = \frac{\sqrt{\text{variance de l'estimation}}}{\text{estimation}}$$

Cette mesure contribue à faciliter l'interprétation quant à la précision d'une estimation. Plus le CV est élevé, moins l'estimation est précise et vice versa. Notamment, il faut faire preuve de prudence lors de l'utilisation des estimations dont le CV est très élevé (supérieur à 25 %), puisque cela indique une très grande variabilité. Ces estimations ne devraient être utilisées qu'à titre indicatif. Le CV a également l'avantage de permettre de comparer la précision de différentes estimations. Le tableau 4 présente la relation entre la valeur du CV et la précision de l'estimation utilisée par l'Institut.

Dans les tableaux produits pour l'Office, la précision est exprimée en CV, soit avec une valeur numérique et une cote. L'Office peut se baser sur les indications présentées dans le tableau ci-dessous dans ses publications pour convertir les CV en « cotes ». Le rapport descriptif produit par l'Institut exprime la précision des estimations avec une cote sous forme d'étoile, tel que cela est présenté dans le tableau 4.

Dans le cas du volet auprès des municipalités et des arrondissements de la Ville de Montréal, où le plan d'échantillonnage était un recensement des unités, il n'y a pas à proprement parler d'erreurs dues à l'échantillonnage. Toutefois, la non-réponse aléatoire est prise en compte. On considère ce type d'erreurs comme des erreurs dues à l'échantillonnage.

La grande majorité des estimations de proportions et de moyennes effectuées pour l'ensemble des unités du volet auprès des établissements ont un CV < 15 %, ce qui signifie qu'elles sont assez précises et qu'elles peuvent être utilisées sans crainte. Toutefois, lorsque l'estimation d'une proportion est très petite, la précision de celle-ci tend à être moins satisfaisante (CV > 15 %). De plus, les cas d'estimations de proportions sur des domaines particuliers – tels que « si oui, passez à la question X » – sont parfois moins précis (CV > 15 %), car il y a moins de répondants admissibles pour répondre à la question. Enfin, le croisement d'une variable d'analyse (par exemple, la question Q2⁸ divisée en quatre catégories) avec une variable contenant également plusieurs catégories (par exemple, la variable du secteur d'activité, divisée en six catégories) va régulièrement engendrer des CV très grands dans les tableaux à cause du petit nombre de répondants sur lesquels sont basées les estimations. La prudence est alors requise dans l'interprétation des résultats. Ces estimations devraient être toujours accompagnées de leur CV exprimé en signe ou de leur intervalle de confiance.

Tableau 4

Relation entre la valeur d'un coefficient de variation et la précision de l'estimation associée

Bornes	Cote Lettres	Cote Étoiles	Mention
CV ≤ 5 %	A		Excellente précision
5 % < CV ≤ 10 %	B		Très bonne précision
10 % < CV ≤ 15 %	C		Bonne précision
15 % < CV ≤ 25 %	D	*	Précision passable ; interpréter avec prudence
CV > 25 %	E	**	Faible précision ; estimation imprécise, fournie à titre indicatif seulement

8. La question Q2 portait sur le temps écoulé depuis la dernière embauche, promotion ou mutation.

5.2 Erreurs non dues à l'échantillonnage

Les erreurs qui ne sont pas dues à l'échantillonnage incluent les différents types de biais pouvant influencer sur les résultats : biais attribuables à la non-réponse, biais liés au mode de collecte, biais de réponse, surcouverture ou sous-couverture de la population par la base de sondage, etc. Les erreurs non dues à l'échantillonnage ne sont pas formellement quantifiables, comme le sont celles de la première catégorie. Il est donc important de tenter le plus possible de détecter leur présence et d'évaluer, le cas échéant, leur ampleur et leur effet.

5.2.1 Taux de réponse global et taux de non-réponse partielle

Le taux de réponse global à une enquête est un indicateur de la qualité de celle-ci, notamment au regard des biais dus à la non-réponse qui pourraient être introduits dans les estimations. En effet, plus le taux de non-réponse est élevé, plus il y a de risques que des biais soient introduits dans les estimations, les non-répondants pouvant présenter des caractéristiques différentes de celles des répondants.

L'Institut, comme tout organisme statistique, veille à limiter les risques de biais dans les enquêtes. C'est pourquoi il cherche à s'assurer du meilleur taux de réponse possible, allant même jusqu'à s'abstenir de procéder à l'inférence statistique lorsque ce taux n'atteint pas un niveau satisfaisant. Dans un tel cas, des analyses de biais sont faites afin de mesurer les risques d'erreur, et des correctifs sont apportés avant la production des estimations.

Dans le cas présent, le taux de réponse global se situe à 83,3 % pour le volet auprès des établissements. Ce taux est un peu plus élevé que l'hypothèse initiale de 80 %. Pour le volet auprès des municipalités et des arrondissements de la Ville de Montréal, le taux de réponse global est de 94,7 %, ce qui est aussi plus élevé que l'hypothèse initiale de 90 %. Toutes les strates de pondération des deux volets ont des taux de réponse pondérés satisfaisants puisqu'ils sont tous plus élevés que 70 %. Notons que la majorité des taux dépasse 80 %. Ces résultats, combinés aux ajustements pour la non-réponse, permettent de dissiper toute inquiétude et de conclure que les risques de biais sont faibles.

Pour ce qui est de la non-réponse partielle, c'est-à-dire de la non-réponse à l'une ou l'autre des questions, il est habituel de considérer qu'un taux de 5 % ou moins ne suscite pas d'inquiétude. Pour le volet auprès des établissements, toutes les questions ont un taux de non-réponse partielle inférieur à 5 %, à l'exception de la Q14 et de la Q24, qui ont un taux de non-réponse partielle de 18,8 % et de 5,7 %, respectivement. Pour le volet auprès des municipalités et des arrondissements de la Ville de Montréal, seule la question Q14 a un taux de non-réponse partielle supérieur ou égal à 5 % (11,4 %).

Notons que pour la Q14, qui porte sur le niveau de compétences recherché, certains chemins du questionnaire Web amenaient la personne répondante à sauter cette question. Ainsi, dans quelques cas, des organisations répondantes ont souhaité une compétence linguistique, mais elles n'ont pas pu préciser le niveau de compétences recherché. Une analyse de biais a été réalisée afin de vérifier si le taux de non-réponse élevé a affecté la qualité de l'estimation. Cela a permis de conclure que, malgré un taux de non-réponse élevé (18,8 %), le risque de biais est demeuré faible.

5.2.2 Réponses aux questions

En général, les questions ont été bien comprises par la majorité des organisations répondantes ; du moins, il n'y a pas eu d'indications contraires. Il n'existe pas de moyen absolu de valider la véracité de toutes les réponses reçues dans le cadre d'une enquête. Cependant, dans le cas présent, rien ne laisse supposer que des organisations répondantes ont délibérément fourni des réponses erronées.

5.2.3 Couverture de la base de sondage du volet auprès des établissements

Même si l'on peut supposer que la base de sondage contient la presque totalité des établissements qui étaient en activité au moment de sa mise à jour (juin 2023 dans le cas présent), il est aussi certain qu'elle compte des établissements qui ont cessé leurs activités. Ce phénomène est connu sous le nom de surcouverture de la population étudiée par la base de sondage.

Cette surcouverture fait qu'il est difficile de connaître la taille exacte de la population étudiée. Bien entendu, on peut faire des hypothèses : dans certaines enquêtes, il est possible de déceler toutes les unités échantillonnées qui ne sont pas admissibles et d'extrapoler ce résultat à l'ensemble de la population étudiée.

Dans la présente enquête, il n'a pas été possible de déterminer le statut de tous les établissements. Dans bien des cas de non-réponse, il n'était pas facile d'attribuer le statut d'admissibilité ou de non-admissibilité à l'enquête. Voici quelques exemples de non-réponse où le statut était ambigu : pas de réponse au numéro composé, mauvais numéro de téléphone, ligne occupée. Il fallait donc faire l'hypothèse suivante : parmi ces cas se trouvent un certain nombre de cas admissibles et un certain nombre de cas inadmissibles. Ces nombres ont pu être estimés, par strate, à l'aide de la répartition des établissements dont le statut était connu. En déterminant un nombre d'établissements non répondants admissibles, il a été possible de donner un poids aux établissements répondants pour que ceux-ci représentent tous les cas admissibles de l'échantillon. De cette façon, et en faisant la somme des poids sur l'ensemble des établissements répondants, une estimation de la population admissible à l'enquête a pu être produite.

6

La confidentialité

L'Institut est tenu, en vertu de la loi, de protéger la confidentialité des renseignements fournis par les organisations répondantes. Les résultats produits sous forme de tableaux sont donc vérifiés. Lors d'un tel examen, il est possible de déceler un type de divulgation, soit une cellule d'un tableau contenant peu d'établissements répondants, soit un petit nombre contribuant à un grand pourcentage d'une estimation présentée dans le tableau.

L'examen des tableaux produits pour la présente enquête a permis de déterminer que le premier type de divulgation pouvait se produire. Les tableaux présentant des problèmes de confidentialité ont été masqués, soit par un regroupement de catégories, soit par un masquage primaire et secondaire des cellules problématiques.

7 L'appréciation globale

Les taux de réponse globaux et par volet ont dépassé les objectifs. En effet, le taux de réponse global obtenu pour le volet auprès des municipalités et des arrondissements est de 94,7 %, tandis qu'il est supérieur à 80 % pour pratiquement toutes les strates du volet auprès des établissements. Dans le contexte général de l'enquête, et en tenant compte des explications présentées dans les sections précédentes, on peut conclure que le potentiel d'analyse des données est très bon et qu'il n'existe pas de problème apparent à inférer les résultats à la population étudiée.

Aussi, comme c'est le cas pour toute enquête, il y a lieu de faire preuve de prudence dans l'usage des résultats qui sont accompagnés d'une forte marge d'erreur ou d'un coefficient de variation supérieur à 15 %.

Enfin, bien que l'imputation n'ait pas été une méthode privilégiée pour contrecarrer la non-réponse partielle, elle a été employée dans des cas très précis. Son incidence demeure négligeable pour des estimations à l'échelle du Québec.

Annexe 1

Population de l'enquête

La population visée par le volet *établissements* de l'*Enquête sur les exigences linguistiques dans les entreprises et les municipalités du Québec en 2023* est constituée des établissements employant cinq personnes et plus dans toutes les industries du Système de classification des industries de l'Amérique du Nord (SCIAN) 2022, à l'exception des industries suivantes :

- SCIAN 22 Services publics
- SCIAN 53111 Bailleurs d'immeubles résidentiels et de logements
- SCIAN 53112 Bailleurs d'immeubles non résidentiels (sauf les mini-entrepôts)
- SCIAN 53119 Bailleurs d'autres biens immobiliers
- SCIAN 533 Bailleurs de biens incorporels non financiers (sauf les œuvres protégées par le droit d'auteur)
- SCIAN 551113 Sociétés de portefeuille
- SCIAN 62 Soins de santé et assistance sociale
- SCIAN 8131 Organismes religieux
- SCIAN 8132 Fondations et organismes de charité
- SCIAN 8133 Organismes d'action sociale
- SCIAN 8134 Organisations civiques et sociales
- SCIAN 81391 Associations de gens d'affaires
- SCIAN 81394 Organisations politiques
- SCIAN 91 Administrations publiques

Dans les tableaux de résultats, les établissements sont regroupés en six secteurs industriels de la façon suivante. Les secteurs dont le nom est accompagné d'une étoile sont exclus de la population visée par l'enquête.

Fabrication et services à forte concentration de connaissances :

- SCIAN 3254 Fabrication de produits pharmaceutiques et de médicaments
- SCIAN 334 Fabrication de produits informatiques et électroniques
- SCIAN 3364 Fabrication de produits aérospatiaux et de leurs pièces
- SCIAN 54 Services professionnels, scientifiques et techniques

Fabrication à concentration moyenne de connaissances :

- SCIAN 21 Extraction minière, exploitation en carrière, et extraction de pétrole et de gaz
- SCIAN 311 Fabrication d'aliments
- SCIAN 312 Fabrication de boissons et de produits du tabac
- SCIAN 313 Usines de textiles
- SCIAN 314 Usines de produits textiles
- SCIAN 322 Fabrication du papier
- SCIAN 323 Impression et activités connexes de soutien
- SCIAN 324 Fabrication de produits du pétrole et du charbon
- SCIAN 325 Fabrication de produits chimiques
 - sauf SCIAN 3254 Fabrication de produits pharmaceutiques et de médicaments
- SCIAN 326 Fabrication de produits en plastique et en caoutchouc
- SCIAN 327 Fabrication de produits minéraux non métalliques
- SCIAN 331 Première transformation des métaux
- SCIAN 332 Fabrication de produits métalliques
- SCIAN 333 Fabrication de machines
- SCIAN 335 Fabrication de matériel, d'appareils et de composants électriques
- SCIAN 336 Fabrication de matériel de transport
 - sauf SCIAN 3364 Fabrication de produits aérospatiaux et de leurs pièces
- SCIAN 339 Activités diverses de fabrication

Services à concentration moyenne de connaissances :

- SCIAN 41 Commerce de gros
- SCIAN 51 Industrie de l'information et industrie culturelle
- SCIAN 52 Finance et assurances
- SCIAN 53 Services immobiliers et services de location et de location à bail
 - sauf SCIAN 53111 Bailleurs d'immeubles résidentiels et de logements*
 - sauf SCIAN 53112 Bailleurs d'immeubles non résidentiels (sauf les mini-entrepôts)*
 - sauf SCIAN 53119 Bailleurs d'autres biens immobiliers*
 - sauf SCIAN 533 Bailleurs de biens incorporels non financiers (sauf les œuvres protégées par le droit d'auteur)*
- SCIAN 55 Gestion de sociétés et d'entreprises
 - sauf SCIAN 551113 Sociétés de portefeuille*
- SCIAN 56 Services administratifs, services de soutien, services de gestion des déchets et services d'assainissement
- SCIAN 61 Services d'enseignement
- SCIAN 71 Arts, spectacles et loisirs

Fabrication à faible concentration de connaissances :

- SCIAN 11 Agriculture, foresterie, pêche et chasse
- SCIAN 315 Fabrication de vêtements
- SCIAN 316 Fabrication de produits en cuir et de produits analogues
- SCIAN 321 Fabrication de produits en bois
- SCIAN 337 Fabrication de meubles et de produits connexes

Services à faible concentration de connaissances :

- SCIAN 44-45 Commerce de détail
- SCIAN 48-49 Transport et entreposage
- SCIAN 72 Services d'hébergement et de restauration
- SCIAN 81 Autres services (sauf les administrations publiques)
 - sauf SCIAN 8131 Organismes religieux*
 - sauf SCIAN 8132 Fondations et organismes de charité*
 - sauf SCIAN 8133 Organismes d'action sociale*
 - sauf SCIAN 8134 Organisations civiques et sociales*
 - sauf SCIAN 81391 Associations de gens d'affaires*
 - sauf SCIAN 81394 Organisations politiques*

Construction

- SCIAN 23 Construction

La population visée par le volet *municipalités* de l'enquête est constituée de toutes les municipalités québécoises de 5 000 personnes ou plus et, en ce qui concerne la Ville de Montréal, de chacun des arrondissements de cette municipalité. Pour ce volet, plusieurs sources de données ont été nécessaires pour la création de la base de sondage. En effet, l'Institut a utilisé deux fichiers du site Internet Données Québec⁹ afin de dresser une liste de 189 municipalités et des 19 arrondissements de la Ville de Montréal.

9. [Données Québec – Données ouvertes du Québec \(donneesquebec.ca\)](https://donneesquebec.ca).

Annexe 2

Résultats de collecte

Tableau 5

Résultats finaux de la collecte des données

Catégorie	Résultats de la collecte	Nombre d'employeurs
Cas inadmissibles		449
	Fermé définitivement ou temporairement	74
	Moins de 5 employé(e)s	197
	Doublon	2
	Secteur d'activité exclu	6
	Inadmissible au questionnaire Web	170
Cas non répondants		503
	Refus	30
	Mauvais numéro ou ligne défectueuse	70
	Incapable de répondre à l'enquête	5
	À rappeler	12
	Répondeur	146
	Pas de réponse	19
	Autres cas (ligne occupée, réexpédition, expédition, etc.)	221
Cas répondants		2 676
	Par la poste ou télécopieur	10
	Par téléphone	125
	Par le Web	2 541
TOTAL		3 628

« Une organisation
statistique performante
au service d'une société
québécoise en évolution »