

**Méthodologie de l'Enquête sur les ventes d'aliments et
de boissons prêts à servir, 2009**

**Michèle Picard-Flibotte
Javier Oyarzun
Eric Pelletier**

**Division des méthodes d'enquêtes auprès des entreprises
Statistique Canada**

21 mai 2010

Table des matières

Table des matières.....	2
1. Objectif de l'enquête.....	3
2. Base de sondage.....	3
3. Échantillonnage et stratification	3
4. Vérification et imputation	6
4.1 Définition d'un questionnaire rempli.....	6
4.2 Convention pour l'entrée des réponses	6
4.3 Critères généraux de sélection des donneurs	6
4.4 Critères de sélection généraux à utiliser dans le cadre de l'imputation.....	7
4.5 Variables dérivées.....	7
4.5.1 Strates.....	7
4.5.2 Région	7
4.5.3 Groupe d'industries.....	8
4.6 Taux d'imputation.....	8
5. Pondération	9
6. Taux de réponse	10

1. Objectif de l'enquête

Cette nouvelle enquête se veut une étude rapide des ventes en deçà de 4 \$ dans l'industrie de la restauration et la vente d'aliments et de boissons prêts à servir en Ontario. Elle prend la forme d'un très court questionnaire (deux questions) et permet d'évaluer les proportions de vente actuellement non taxable. Cela est fait dans le but d'éventuellement coordonner le partage d'une taxe de vente harmonisée dans cette province.

2. Base de sondage

Pour créer la base de sondage, la Registre des entreprises a été utilisé. L'unité statistique est l'établissement.

3. Échantillonnage et stratification

Pour sélectionner l'échantillon, les paramètres suivants ont été considérés :

Taille visée de 1500 unités

Taux de réponse espéré de 60 %

CVs visés par domaine

- 10 % pour les groupes d'industrie 6, 11 et 12

- 5,8 % pour les autres groupes d'industrie

Les domaines sont les suivants :

Table 1 : SCIAN inclus dans chacun des domaines

Domaines : Groupe d'industries	Codes du SCIAN inclus
1- Boulangeries	SCIAN 311811 et 445291
2- Cafétérias; cafés-restaurants; casse-croûte et snack-bars; beigneries; restaurants à service restreint	SCIAN 722210
3- Friteries mobiles; stands à hot-dogs; stands à café; cantines et comptoirs mobiles	SCIAN 722330
4- Salons-bars, tavernes et bars	SCIAN 722410
5- Dépanneurs	SCIAN 445120 et 447110
6- Centres des congrès	SCIAN 531120
7- Entrepreneurs en restauration	SCIAN 722310
8- Restaurants à service complet	SCIAN 722110
9- Épicerie	SCIAN 445110
10- Hôtels et motels	SCIAN 7211
11- Maisons de chambres	SCIAN 7213
12 – Clubs privés, amicales et filiales de la Légion	SCIAN 813410
13- Casse-croûte	SCIAN 512130, 7111, 7112, 71131, 712, 7131, 713210, 713299, 7139
14- Distributeurs automatiques*	SCIAN 454210

* Seulement les opérateurs de distributeurs automatiques qui ont vendu des repas ont été incluse sur la base de sondage.

Chacun des domaines a été stratifié selon leur taille basée sur le revenu pour obtenir une strate à tirage nul, 2 strates à tirage partiel et une strate à tirage complet. Toutefois, comme c'est la première occasion de cette enquête et que la relation entre les ventes d'aliments et de boissons prêts à servir sous 4 \$ et les ventes totales d'aliments et de boissons prêts à servir est inconnue, un échantillon a été sélectionné parmi les strates à tirage nul. Par la suite, le système généralisé d'échantillonnage a été utilisé pour faire l'allocation selon la méthode de minimisation du coût total et la sélection de l'échantillon.

Tableau 2 : Taille de la population et de l'échantillon par NAICS et par strate

GroupeInt	SCIAN(SCIAN)	Sizeclassh								Total	
		0		1		2		3			
		Tirage nul		Tirage partiel petit		Tirage partiel moyen		Tirage complet		N	n
		N	n	N	n	N	n	N	n		
	311811	437	39	2	0	52	10	11	11	502	60
	445291	285	19	151	12	45	10	10	10	491	51
1	Total	722	58	153	12	97	20	21	21	993	111

2	722210	5840	30	5767	58	1099	52	18	18	12724	158
3	722330	283	13	98	8	23	7	7	7	411	35
4	722410	540	12	441	20	178	22	18	18	1177	72
	445120	1741	19	3265	32	145	5	4	4	5155	60
	447110	912	8	49	1	481	8	19	19	1461	36
5	Total	2653	27	3314	33	626	13	23	23	6616	96
6	531120	17223	87	5048	52	488	50	47	47	22806	236
7	722310	667	8	0	0	238	8	13	13	918	29
8	722110	5615	28	5097	52	1453	52	10	10	12175	142
9	445110	2297	23	713	33	271	35	6	6	3287	97
	721111	537	7	3	0	201	7	48	48	789	62
	721112	69	1	94	0	18	0	0	0	181	1
	721113	160	2	73	2	32	1	7	7	272	12
	721114	339	2	544	7	8	0	0	0	891	9
	721191	432	5	27	0	0	0	0	0	459	5
	721192	410	5	83	1	2	0	0	0	495	6
	721198	111	0	26	0	7	0	0	0	144	0
10	Total	2058	22	850	10	268	8	55	55	3231	95
11	721310	37	7	51	7	21	7	9	9	118	30
12	813410	770	8	1532	25	316	33	55	55	2673	121
	512130	122	1	0	0	108	7	0	0	230	8
	711111	194	4	210	2	8	0	2	2	414	8
	711112	56	2	9	0	8	0	2	2	75	4
	711120	31	0	37	0	1	0	0	0	69	0
	711130	484	5	604	6	17	3	0	0	1105	14
	711190	17	1	27	1	0	0	0	0	44	2
	711211	182	2	5	0	19	1	3	3	209	6
	711213	668	4	303	3	29	0	7	7	1007	14
	711218	215	0	13	1	3	0	0	0	231	1
	711311	106	0	23	1	7	0	2	2	138	3
	711319	79	0	24	0	9	0	1	1	113	1
	712111	36	1	30	1	3	0	1	1	70	3
	712115	3	0	8	0	3	0	2	2	16	2
	712119	21	1	32	0	8	0	1	1	62	2
	712120	6	0	13	0	1	0	0	0	20	0
	712130	24	1	7	0	2	0	2	2	35	3
	712190	22	0	18	0	5	1	1	1	46	2
	713110	51	1	6	0	7	1	5	5	69	7
	713120	114	0	11	0	6	1	0	0	131	1
	713210	2	0	2	0	2	0	4	4	10	4
	713299	72	1	27	0	31	2	2	2	132	5
	713910	349	7	255	4	217	21	4	4	825	36
	713920	35	0	5	0	17	1	0	0	57	1
	713930	136	0	229	5	29	1	0	0	394	6
	713940	917	11	599	9	113	11	4	4	1633	35
	713950	66	0	137	3	10	1	0	0	213	4
	713990	820	6	754	11	51	6	2	2	1627	25
13	Total	4828	48	3388	47	714	57	45	45	8975	197
14	454210	386	7	18	18	97	14	28	28	529	67
	Total	43919	378	26470	375	5889	378	355	355	76633	1486

4. Vérification et imputation

La présente section rend compte des spécifications pour la vérification et l'imputation. Les tâches (vérification et imputation) sont énumérées selon leur ordre d'exécution. Des instructions détaillées sont également fournies pour la création d'une série de variables dérivées. Ces dernières constituent un élément clé de la stratégie d'imputation et, par conséquent, elles doivent être créées au tout début du processus.

Les règles de vérification permettent de repérer et de corriger les enregistrements présentant des réponses incohérentes, incomplètes ou invalides. S'il n'est pas possible de corriger directement les erreurs, les valeurs inacceptables seront remplacées par les valeurs imputées. Tous les éléments ayant fait l'objet d'une vérification seront marqués d'une puce ronde.

L'imputation indique comment seront corrigés les enregistrements qui ne l'avaient pas déjà été à l'étape de la vérification. En règle générale, il a été déterminé qu'aucune méthode d'imputation susceptible d'introduire un biais ne serait utilisée. Ainsi, l'imputation du donneur se fera d'après les critères de sélection généraux énoncés ci-dessous, seule ou dans le contexte d'une recherche par valeur fixe ou par plus proche voisin en utilisant une variable différente.

La vérification et l'imputation ne peuvent se faire qu'avec les questionnaires qui ont été remplis.

4.1 Définition d'un questionnaire rempli

Un questionnaire est considéré comme rempli si une réponse a été donnée à l'une des questions mentionnées dans le présent document.

4.2 Convention pour l'entrée des réponses

Le questionnaire contient deux questions auxquelles on doit fournir une réponse à l'intérieur d'une fourchette acceptable. Voici comment elles se présentent :

C0300) Veuillez déclarer en dollars (\$) le total de vos ventes d'aliments et de boissons non alcoolisées en 2009. Votre réponse doit être comprise entre 0 \$ et 999 999 999 997 \$.

Pour la deuxième question, le répondant a la possibilité de répondre à la question 1a ou à la question 1b.

C0401) De ce total, veuillez déclarer le pourcentage (%) de ventes exemptes de TVP en 2009. Votre réponse doit être comprise entre 0 % et 100 %.

C0400) De ce total, veuillez indiquer, en dollars, le montant de ventes exemptes de TVP en 2009. Votre réponse doit être comprise entre 0 \$ et 999 999 999 997 \$.

4.3 Critères généraux de sélection des donneurs

Ce document fait souvent référence aux « critères de sélection généraux ». Il s'agit en fait de la méthode générale à appliquer pour sélectionner le donneur pendant le processus d'imputation. Les donneurs sont sélectionnés selon des critères classés par ordre d'importance. Le premier rang correspond au donneur idéal. Vient ensuite le deuxième rang, et ainsi de suite. Chaque rang comporte deux niveaux de correspondance fondés sur les données détaillées de l'industrie.

4.4 Critères de sélection généraux à utiliser dans le cadre de l'imputation

A. Groupe d'industries x strates

B. Groupe d'industries

4.5 Variables dérivées

Une série de variables dérivées sera créée pour les besoins de la mise en œuvre de la stratégie d'imputation, pour la création de tableaux statistiques et pour les chercheurs et analystes à qui ces variables pourraient être utiles.

4.5.1 Strates

Cette variable de taille des revenus est fondée sur la variable revenue dérivée de l'information du Registre des entreprises. Une strate utilisée à des fins d'échantillonnage a été développée à partir de cette information.

Strate	Critères
Tirage complet	La strate à tirage complet se compose des plus grandes entreprises.
Tirage partiel grand	La strate à tirage partiel grand est une sous-strate composée des moyennes entreprises ayant une fraction de sondage plus élevée que la sous-strate des plus petites entreprises à tirage partiel.
Tirage partiel petit	La strate à tirage partiel petit est une sous-strate composée des moyennes entreprises ayant une fraction de sondage plus petite que la sous-strate des plus grandes entreprises à tirage partiel.
Tirage nul	La strate à tirage nul est composée des plus petites entreprises et représente la fraction de sondage la plus faible.

4.5.2 Région

Région	Critères
Onta	Ontario

4.5.3 Groupe d'industries

Groupe d'industries	Critères
Boulangeries	SCIAN 311811 et 445291
Cafétérias; cafés-restaurants; casse-croûte et snack-bars; beigneries; restaurants à service restreint	SCIAN 722210
Friteries mobiles; stands à hot-dogs; stands à café; cantines et comptoirs mobiles	SCIAN 722330
Salons-bars, tavernes et bars	SCIAN 722410
Dépanneurs	SCIAN 445120 et 447110
Centres des congrès	SCIAN 531120
Entrepreneurs en restauration	SCIAN 722310
Restaurants à service complet	SCIAN 722110
Épiceries	SCIAN 445110
Hôtels et motels	SCIAN 7211
Maisons de chambres	SCIAN 7213
Clubs privés, amicales et filiales de la Légion	SCIAN 813410
Casse-croûte	SCIAN 512130, 7111, 7112, 71131, 712, 7131, 713210, 713299, 7139
Distributeurs automatiques*	SCIAN 454210

4.6 Taux d'imputation

Tableau 3: Taux d'imputation pondéré par question et par groupe d'industries

Groupes du SCIAN		Pondéré		
		C0300	C0401	C0400
	Total – Ontario	9.1%	9.8%	9.8%
1	Boulangeries	5.8%	5.6%	5.6%
2	Cafétérias; cafés-restaurants; casse-croûte et snack-bars; beigneries; restaurants à service restreint	14.3%	19.4%	19.4%
3	Friteries mobiles; stands à hot-dogs; stands à café; cantines et comptoirs mobiles	0.0%	0.0%	0.0%
4	Salons-bars, tavernes et bars	0.0%	13.8%	13.8%
5	Dépanneurs	22.2%	27.4%	27.4%
6	Centres des congrès	1.9%	0.0%	0.0%
7	Entrepreneurs en restauration	0.3%	0.6%	0.6%
8	Restaurants à service complet	6.7%	6.4%	6.4%
9	Épiceries	40.5%	16.0%	16.0%
10	Hôtels et motels	0.0%	6.0%	6.0%
11	Maisons de chambres	1.6%	0.0%	0.0%
12	Clubs privés, amicales et filiales de la Légion	4.0%	5.3%	5.3%
13	Casse-croûte	6.1%	8.1%	8.1%
14	Distributeurs automatiques	0.0%	1.2%	1.2%

Tableau 4: Taux d'imputation non pondéré par question et par groupe d'industries

Groupes du SCIAN		Non Pondéré		
		C0300	C0401	C0400
	Total – Ontario	6.2%	11.4%	11.4%
1	Boulangeries	7.0%	5.3%	5.3%
2	Cafétérias; cafés-restaurants; casse-croûte et snack-bars; beigneries; restaurants à service restreint	14.6%	16.7%	16.7%
3	Friteries mobiles; stands à hot-dogs; stands à café; cantines et comptoirs mobiles	0.0%	0.0%	0.0%
4	Salons-bars, tavernes et bars	0.0%	20.0%	20.0%
5	Dépanneurs	17.3%	30.8%	30.8%
6	Centres des congrès	0.8%	0.8%	0.8%
7	Entrepreneurs en restauration	8.3%	16.7%	16.7%
8	Restaurants à service complet	5.3%	9.3%	9.3%
9	Épiceries	28.2%	30.8%	30.8%
10	Hôtels et motels	0.0%	16.7%	16.7%
11	Maisons de chambres	5.9%	0.0%	0.0%
12	Clubs privés, amicales et filiales de la Légion	1.2%	6.0%	6.0%
13	Casse-croûte	5.2%	10.4%	10.4%
14	Distributeurs automatiques	0.0%	6.8%	6.8%

Les taux furent obtenus en utilisant les formules suivantes :

$$TI_{non-pondéré} = (ci)/nj$$

où ci := cellules imputées (si la cellule est imputée, alors ci = 1, sinon ci =0) pour répondant i à la question j, nj := nombre de cellules répondues pour la question j

$$TI_{pondéré} = (ci_w)/nj_w$$

où ci_w := cellules imputées (si la cellule est imputée, alors ci = poids initial, sinon ci =0) pour répondant i à la question j, nj_w:= somme des poids initiaux pour les cellules répondues pour la question j

5. Pondération

Pour le traitement de la non-réponse totale, nous avons eu recours à la re-pondération pour compenser cette non-réponse.

Correction pour la non-réponse

Les unités sélectionnées lors de l'échantillonnage furent regroupées en quatre catégories comme suit :

Tableau 5 : Code de réponse et leur fréquence dans l'échantillon

Code de statut regroupé	Fréquence
NR : non répondant	465
OOS : hors champ	68
OOB : mort	120
REP (INSCOPE) : répondant dans le champ de l'enquête	833

Le poids des non répondants est redistribué entre les répondants qui sont dans le champ de l'enquête, les OOS et les OOB en utilisant les strates du plan d'échantillonnage comme classe de pondération. Habituellement, lorsqu'il n'y a pas suffisamment de répondants dans une strate donnée, un regroupement de strates est nécessaire. Dans cette enquête, seulement deux strates n'avaient qu'un répondant, dont une avec aucun non répondant et l'autre avec un total de ventes d'aliments et de boissons prêts à servir raisonnable. Pour ces raisons, aucune strate n'a été regroupée.

6. Taux de réponse

Pour le calcul du taux de réponse, nous avons calculé les taux de réponse non-pondéré et pondéré. Le taux de réponse non-pondéré calculé à l'estimation est de 56 % tandis que le taux de réponse pondéré est de 60 %. Le taux de réponse non-pondéré permet d'avoir une idée du taux de réponse au niveau de l'échantillon, souvent utilisé lors de la collecte, alors que celui pondéré permet d'avoir une idée globale, c'est-à-dire au niveau de la population.

Les taux ont été obtenus en utilisant les formules suivantes :

$$a) TR_{non-pondéré} = (c + p)/n = 833/1486 = 56,1\%$$

où c := questionnaires complétés, p := questionnaires partiellement complétés, n := taille de l'échantillon

$$b) \text{ pondéré : } TR_{pondéré} = \frac{(c_w + p_w)}{(n_w - oos_w - oob_w - dup_w)} = \frac{36498}{(76633 - 5795 - 9945)} = 59,9\%$$

où c_w := questionnaires complétés pondérés, p_w := questionnaires partiellement complétés pondérés, n_w := taille de l'échantillon pondérée, oos_w := out-of-scope pondérés, oob_w := out-of-business pondérés, dup_w := duplicats pondérés.

Tableau 6 : Taux de réponse non-pondérés par groupe d'industries

Groupes du SCIAN			pop	sample	RES	txrep_nonpondere
	Total – Ontario		76633	1486	833	56.1%
1	Boulangeries	311811 et 445291	993	111	57	51.4%
2	Cafétérias; cafés-restaurants; casse-croûte et snack-bars; beigneries; restaurants à service restreint	722210	12724	158	96	60.8%
3	Friteries mobiles; stands à hot-dogs; stands à café; cantines et comptoirs mobiles	722330	411	35	15	42.9%
4	Salons-bars, tavernes et bars	722410	1177	72	35	48.6%
5	Dépanneurs	445120 et 447110	6616	96	52	54.2%
6	Centres des congrès	531120	22806	236	126	53.4%
7	Entrepreneurs en restauration	722310	918	29	12	41.4%
8	Restaurants à service complet	722110	12175	142	75	52.8%
9	Épiceries	445110	3287	97	39	40.2%
10	Hôtels et motels	721111, 721112, 721113, 721191, 721191, 721192 et 721198	3231	95	66	69.5%
11	Maisons de chambres	721310	118	30	17	56.7%
12	Clubs privés, amicales et filiales de la Légion	813410	2673	121	84	69.4%
13	Casse-croûte	512130, 7111, 7112, 71131, 712, 7131, 713210, 713299 et 7139	8975	197	115	58.4%
14	Distributeurs automatiques	454210	529	67	44	65.7%